Three Essays on Bank Lending

Dissertation

submitted to the
Faculty of Business, Economics and Informatics
of the University of Zurich

to obtain the degree of
Doktorin der Wirtschaftswissenschaften, Dr. oec.
(corresponds to Doctor of Philosophy, PhD)

presented by
Fulvia Fringuellotti
from Italy

approved in July 2019 at the request of

Prof. Dr. Steven Ongena
Prof. Dr. Christoph Basten

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 17.07.2019

The Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Acknowledgements

First and foremost I would like to express my sincere gratitude to my supervisor Steven Ongena for his valuable advice, his kind availability and his continuous support throughout my PhD studies. I am particularly thankful for encouraging me to aim for high quality research and for giving me the opportunity to spend several visiting periods at Bank of Italy to work on unique data. It was a privilege for me to work with him and to draw on his exceptional scientific knowledge to develop my skills as researcher. The fruitful working environment he provided and his excellent guidance helped me to achieve my academic and professional goals.

I also would like to thank my coauthors Ugo Albertazzi, Nicola Branzoli and Manthos Delis for the stimulating discussions, the exciting work conducted together and their great support. Our collaborations allowed me to expand my expertise and to improve my skills as researcher. A special thanks goes to Ugo Albertazzi, Nicola Branzoli and their colleagues at Bank of Italy (in particular, Francesco Columba, Giorgio Gobbi, Stefano Neri and Stefano Siviero) for giving me the chance to visit the Financial Stability Directorate and the Economic Outlook and Monetary Policy Directorate of Bank of Italy to work on two research projects presented in this thesis. This experience was an extraordinary opportunity to be exposed to relevant policy issues, to connect with a highly qualified community of economists and to pursue innovative research.

Furthermore, I am grateful to Christoph Basten for being part of my committee, as well as for the inspiring discussions, the exceptional feedbacks and the constructive suggestions.

I also would like to thank Michel Habib for his advice, his continued support and the excellent working atmosphere at his chair.

Moreover, I want to thank my family and my boyfriend Kevin for their love, patience and great help in going through the doctoral journey. I am also grateful to my colleagues at the Institute for the friendly environment, their prompt help and the pleasant moments shared together.

Additionally, I am grateful for financial support from the European Union Seventh Framework Programme (FP7/2007-2013) under the ERC grant agreement 249415-RMAC. I also would like to thank Knut Are Aastveit, Ida Wolden Bache, Tobias Berg, Emilia

Bonaccorsi di Patti, Alessandro Borin, Margherita Bottero, Martin Brown, Elena Carletti, Geraldo Cerqueiro, Marc Chesney, Nicola Cetorelli, Jerry Coakley, Francesco Columba, Ricardo Correa, Francisco Covas, Olivier De Jonghe, Hans Degryse, Yota Deli, Swati Dhingra, Robin Döttling, Alberto Felettigh, Sigurd Galaasen, Leonardo Gambacorta, Emilia Garcia-Appendini, Pedro Gete, Mariassunta Giannetti, Delis Grombe, Giovanni Guazzarotti, Torbjørn Hægeland, Iftekhar Hasan, Jens Hagendorff, Fédérick Holm-Hadulla, Li He, Artashes Karapetyan, Sotirios Kokas, Alexandros Kontonikas, Anna Kovner, Thomas Lambert, Michael Lamla, Luc Laeven, Seung Lee, Juan-Miguel Londono-Yarce, Mancy Luo, Sai Ma, Kaveh Majlesi, Kevin Meyer, Camelia Minoiu, Stefano Neri, Per Östberg, Georgios Panos, Fabio Parlapiano, Anna Pavlova, Christoph Perignon, Diane Pierret, Andrea Presbitero, Simon Price, Ricardo Reis, Lucia Rizzica, John Rogers, Kasper Roszbach, Farzad Saidi, Joao Santos, Alessandro Secchi, Anatoli Seguravelez, Enrico Sette, Joel Shapiro, Federico Maria Signoretti, Stefano Siviero, Roberto Steri, Daniel Streitz, Javier Suarez, Marta Szymanowska, Leif Anders Thorsrud, Mathijs van Dijkm, Ignazio Visco, Guillaume Vuillemey, Alexander F. Wagner, Wolf Wagner, Wojciech Zurowski and seminar participants at Adam Smith Business School (University of Glasgow), Athens University of Economics and Business, Bank of Italy, Essex Business School (University of Essex), Federal Reserve Bank of New York, Federal Reserve Board, Montpellier Business School, Norges Bank, Rotterdam School of Management, University of Lausanne, University of Zurich, and participants at the Banking Research Network Workshop at Bank of Italy, the CEPR's Endless Summer Conference on Financial Intermediation and Corporate Finance, the SFI Research Days 2017, and the 11th Swiss Winter Conference on Financial Intermediation for very helpful comments and suggestions.

Thank you!

Fulvia Fringuellotti, Zurich, April 2019

Thesis Overview

This thesis studies the connections between various facets of bank lending and the real economy. The focus is on three main issues: the real effects of bank monitoring, the allocation of interest-rate risk in the mortgage market, and the impact of access to credit on income inequality.

The historical evidence, and particularly the recent financial crisis, show that banking crises produce huge negative externalities on the financial system, on public finances and on the economy as a whole. In light of that, fostering a safe and sound banking system is of crucial importance. This requires to study the risk-taking incentives of banks and their implications for financial stability.

In the activity of lending, banks are typically exposed to credit risk. The exposure to credit risk depends on two crucial factors: the selection of borrowers (ex ante screening), and the extent to which banks monitor their borrowers (ex post monitoring). Bank monitoring consists in all supervising activities aimed at verifying and improving the likelihood that a borrower repays his debt. Since the early banking literature, monitoring has been identified as a major factor explaining the existence of banks. Nevertheless, from an empirical perspective, the real effects of bank monitoring are rather unexplored.

In the first paper (Chapter 1) titled "The Effect of Bank Monitoring on Loan Repayment" my coauthor and I investigate bank monitoring and its effect on the likelihood of loan repayment. In particular, we test empirically if monitoring is able to reduce delinquency rates. To this end, we exploit data from the Italian Credit Register of Bank of Italy, which contains loan-level information on virtually all loans extended in Italy. This allows us to build a novel proxy for bank monitoring, which is based on the requests for information to the Credit Register made by banks on their existing borrowers. This proxy is able to directly capture the effort exerted by loan officers in carrying out the various tasks that are useful to monitor their borrowers. From a methodological perspective, we are able to investigate the causal effect of monitoring on loan repayment using taxation as a source of exogenous variation in bank monitoring. This approach is backed by a theoretical model that we develop to show how a corporate income tax affects bank monitoring incentives.

Our main findings suggest that the activity of monitoring improves bank stability by reducing delinquency rates in a significant way. This has important implications for financial stability, as it suggests that lenders that extend credit without monitoring their borrowers may experience higher default rates. Additional results provide relevant insights on what are the drivers of bank incentives to monitor borrowers. Particularly relevant is the evidence that tax policy is able to affect these incentives in a substantial way, implying that taxation can be used as a valid tool for financial stability.

The recent financial crisis highlighted also the importance of mortgage credit for macroeconomic stability. In the second paper (Chapter 2) "Fixed Rate versus Adjustable Rate Mortgages: Evidence from Euro Area Banks" my coauthors and I focus on a specific dimension of this issue. In particular, a striking feature of the mortgage market in the euro area is the very large heterogeneity across countries in the granting of fixed versus adjustable rate mortgages. This has two major implications. First, the allocation of interest-rate risk between the banking sector and the real sector differ across countries, with direct consequences for financial stability. Second, the transmission of monetary policy is heterogeneous across countries, posing relevant concerns for policy makers. Typically mortgages represent a major liability in the balance sheet of most households. Therefore, in systems where adjustable rate mortgages are dominant, households are particularly exposed to interest-rate risk. In light of that, understanding the reasons why residential mortgages carry a fixed or an adjustable interest rate is crucial in order to derive normative insights.

With this caveat in mind, we investigate the determinants of mortgage choice in the euro area. To this end, we exploit bank-level information on the lending activity of a representative sample of banks in the euro area. In particular, we examine to what degree the wide cross-country heterogeneity in the prevalent interest rate of housing loans is driven by borrower and bank characteristics. To disentangle these two components we compare the lending patterns of cross-border banks, which operate in more than one country.

Our results suggest that borrower characteristics at the country level play a prominent role. These include, the level of financial literacy, the macroeconomic history of a country, and the suitability of local mortgages to back covered bonds and mortgage-backed securities. Particularly interesting is the result that borrowers tend to select the type of mortgage that allows a higher degree of consumption smoothing during periods of economic downturn. In general, our findings suggest that the large heterogeneity in the prevalent type of mortgage essentially reflects an optimal allocation of interest-rate risk, which differs across countries given the asynchronous business cycles and the diverse institutional environments. Therefore, in terms of policy implications our analysis reveals that it is not advisable to influence the mortgage market pressing banks to take on more duration risk in contexts where adjustable rate mortgages prevail.

A further legacy of the recent financial crisis is the strengthening in the pre-existing phenomenon of increasing income inequality. In fact, in the last decades the gap between the rich and the poor has risen in most OECD countries, posing serious concerns for social cohesion and economic growth. This moved income inequality at the top of the policy and research agenda, yielding a lively debate on the potential tools to tackle this trend. Access to credit is likely to play a prominent role. In fact, the academic literature identifies a strong link between financial depth and income inequality at the macro level.

In the third paper (Chapter 3) "Credit and Income" my coauthors and I investigate the effect of bank credit origination on individuals' income. The goal of this study is to provide relevant insights on the finance-inequality nexus from a micro perspective. To this end, we use a unique data set of business loan applications to a single European bank. Our focus is on loan applications from small and micro enterprises that are majority-owned by individuals who have an exclusive relationship with the bank. For these applicants, the bank has information on the income of the business owner and decides whether to grant the loan on the basis of a cutoff rule. Specifically, each applicant is given a credit score at the time of the loan application. Credit is granted to applicants whose credit score is above the cutoff, and denied otherwise. Our methodological approach builds on the idea that individuals whose credit score is around the cutoff are virtually the same in terms of credit quality. Thus, to estimate the effect of credit origination on individual's income, we compare the level of income of accepted and rejected applicants after the loan decision has taken place.

Our findings reveal that a loan origination increases the recipient's income in a significant way. The effect is stronger in low-income and rural areas compared to high-income and urban areas, as well as during periods of crisis compared to normal times. These results have important implications for income inequality, pointing to a negative finance-inequality nexus. Overall, our findings provide empirical support to policy interventions aimed at increasing access to credit to reduce income inequality, such as those of the European Bank for Reconstruction and Development (EBRD) and the European Investment Bank (EIB).

Contents

Chapter 1

The Effect of Bank Monitoring on Loan Repayment

We investigate the effect of bank monitoring on loan repayment. Using granular loan-level information from the Italian Credit Register, we build a novel measure of bank monitoring, which is based on bank requests for information on their existing borrowers. We perform a causal analysis, exploiting the Italy Regional Production Tax, IRAP, as a source of exogenus variation in bank monitoring. Our approach is supported by a theoretical model predicting that a decrease in the tax rate improves bank incentives to monitor borrowers. We find that an increase in the number of requests for information, as driven by a 1 percentage point decrease in the IRAP tax rate, reduces the probability of loan distress by almost 4 percentage points two quarters ahead.

Branzoli, N., Fringuellotti, F., 2019. The Effect of Bank Monitoring on Loan Repayment. Working Paper.

1.1 Introduction[1]

Bank monitoring consists in all supervising activities aimed at verifying and improving the likelihood that a borrower complies with its loan obligations (Townsend, 1979; Diamond, 1984; Gale and Hellwig, 1985; Krasa and Villamil, 1992; Holmstrom and Tirole, 1997). Since the early banking literature, monitoring has been identified as a major factor explaining the existence of banks (Diamond, 1984). Yet the real effects of bank monitoring are rather unexplored.

In this paper we investigate how bank monitoring affects loan repayment. From an empirical perspective, assessing this causal relation is challenging. First, it is difficult to directly measure the intensity of bank monitoring, as it is usually not observable. Second, the repayment performance of the borrower is likely to influence bank monitoring as well, implying that causality can go in either direction.

We use quarterly loan-level information from the Italian Credit Register to build a novel proxy for bank monitoring. This proxy is based on the requests for information to the Credit Register made by banks on their existing borrowers. These requests allow to get information on the amount of each loan granted by other banks to a specific borrower, as well as on the objective conditions of deterioration of any of the individual exposures. We interpret a request for information as evidence that the bank decides to take a closer look at one of its borrowers. Hence, we use the number of requests for information occurring in a given quarter as our proxy for bank monitoring. We acknowledge that a bank can monitor its borrowers also in other ways, for example by talking to the firm's managers or analyzing the borrower's financial report. Our proxy is not intended to quantify all these activities, but rather to capture an observable evidence of the effort exerted by banks in carrying out the various tasks that are useful to monitor their borrowers. With this caveat in mind, we analyze the appropriateness of our proxy for bank monitoring. As expected, we find that banks monitor more intensively opaque borrowers (e.g. small firms with a shorter credit relationship with the bank), riskier borrowers (those with lower credit rating) and in periods of economic downturns, when the risk of credit deterioration increases.

A preliminary analysis shows that the number of requests for information is negatively

[1] The opinions expressed in this paper are those of the authors only and do not necessarily reflect those of the Bank of Italy.

related to the future probability of loan distress, suggesting that bank monitoring may have a positive effect on loan repayment. However, this evidence is not enough to establish causation. To investigate the causal effect of bank monitoring on loan repayment we use taxation as a source of exogenous variation in bank monitoring. Our approach builds on a theoretical model that we develop to describe the effects of a corporate income tax on bank incentives to monitor borrowers. In this model a representative bank determines the optimal monitoring effort, capital ratio and lending rate by maximizing its expected profits. An increase in the corporate tax rate implies a decrease in net profits after tax and a reduction in the capital ratio, which are only partially counteracted by an increase in the lending rate. Overall, this means that an increase in the corporate tax rate results in lower expected profits. Intuitively, monitoring incentives are stronger the higher is the fraction of expected profits to shareholders. In fact, the model predicts that bank monitoring increases when the corporate tax rate decreases. We therefore rely on this prediction to outline the identification strategy of our empirical study.

We address our research question by exploiting exogenous variation in bank monitoring driven by the Italy Regional Production Tax ("Imposta Regionale Attivitá Produttive", IRAP) rate applied to banks. This tax rate is set at the regional level and varies both across regions and over time. We focus our attention on small banks operating at the regional level, which represent the financial intermediaries mostly affected by changes in this local tax.

As a first step, we consider a 2SLS model in which we estimate the effect of bank monitoring on the likelihood that the loan is nonperforming two quarters ahead. To this end, bank monitoring is instrumented with the IRAP tax rate. A major advantage of our dataset is that it includes a sizable number of borrowers having multiple credit relationships. Thus, we rely on time-varying borrower fixed effects to control for any observable and unobservable borrower's condition that may affect the ability to meet the contractual obligations. This implies that we estimate the effect of interest by comparing the repayment performance of different loans granted by different banks to the same firm.

Consistently with our theoretical prediction, we find that an increase in the IRAP tax rate implies a decrease in bank monitoring. More importantly, we find that monitoring has a positive and statistically significant effect on loan repayment. The economic magnitude is substantial: an increase in the number of requests for information, that corresponds to a 1 percentage point decrease in the IRAP tax rate, reduces the probability that the loan becomes nonperforming by 3.9 percentage points two quarters ahead. This result is big in economic terms if we take into account that the probability of loan distress is roughly 11% in our sample.[2]

The main limitation of our proxy for bank monitoring is that it captures only a fraction of the entire monitoring activity exerted by a bank. In fact, a bank can monitor its borrower also in other ways, for example by checking the company's financial report, by visiting the firm on site, and by providing advisory services. Because variation in the IRAP tax rate is likely to affect bank incentives with respect to any potential monitoring approach, we are able to extend our analysis investigating the effect of the overall intensity of bank monitoring on loan repayment.

To this end, we estimate a reduced form model in which we directly use the IRAP tax rate to capture the entire monitoring activity of the bank. Our results suggest that a 1 percentage point decrease in the IRAP tax rate leads to a reduction in the probability of loan distress by 4.9 percentage points. The magnitude of this effect is close and somewhat higher than what detected in the 2SLS model. Despite the two models are not directly comparable, as they are estimated using a slightly different sample, this result suggests that the requests for information are highly correlated with other forms of bank monitoring. This means that our novel variable is able to capture to a large extent the overall effect of bank monitoring on loan repayment.

Our paper relates to the literature studying the role of banks as delegated monitors (Diamond, 1984; Krasa and Villamil, 1992; Holmstrom and Tirole, 1997). We contribute to this literature in various ways. First, we provide a novel and direct proxy for bank monitoring at the loan level. Our approach to gauge bank monitoring is in the same vein of Gustafson et al.(2017), but we use bank requests for information on lending exposures rather than on financial statements. More importantly, and to the best of our knowledge, we are the first to investigate the causal effect of bank monitoring on the likelihood of loan repayment. We show empirical evidence that bank monitoring is valuable, as it improves in a substantial way the borrower's repayment performance. Our results provide useful insights for regulators and policy makers, especially in light of the topical debate on the

[2]This figure refers to the sample including only firms having multiple bank relationships that we use in our main regression analysis.

"originate-to-distribute" model (Brunnermeier, 2009). Our findings suggest that lenders that extend credit without monitoring their borrowers may experience higher default rates, posing relevant concerns for financial stability.

This work also contributes to the strand of literature investigating the relation between bank stability and risk-taking. This literature highlights that an increase in the survival probability of the bank and, hence, in the likelihood of retaining rents from lending, weakens bank incentives to take on risk (Allen et al., 2011; Mehran and Thakor, 2011; Dell'Ariccia et al., 2014; Jiménez et al., 2014; Bhat and Desai, 2017; Dell'Ariccia et al. 2017; Jiménez et al., 2017). Typically, the relation between bank stability and risk taking has been analysed by focusing on variation in the capital ratio (Allen et al., 2011; Mehran and Thakor, 2011; Dell'Ariccia et al., 2014; Bhat and Desai, 2017; Jiménez et al., 2017) or in the level of the policy rate (Jiménez et al., 2014; Dell'Ariccia et al., 2017). We contribute to this literature by documenting, both theoretically and empirically, that higher bank stability, as driven by a decrease in the corporate tax rate, results in higher monitoring incentives. Our result is consistent with the existing literature, and, specifically, with the idea that bank shareholders expecting higher profits have more "skin in the game" and are less inclined to take on risk.

Finally, this work also complements the literature on the effects of taxation on bank risk-taking. Exploiting the same variation in the IRAP tax rates used in this work and performing an analysis at the bank-level, Gambacorta et al. (2017) show that an increase in the corporate tax rate leads to a riskier composition of the asset side of bank's balance sheet. The model developed in this paper provides a theoretical explanation for their result. Moreover, looking at a diverse set of tax interventions, Schepens (2016), Devereux et al. (2015), Carletti et al. (2018), and Célérier et al. (2018) document that taxation is able to shape the riskiness of bank assets in a significant way. We contribute to this literature showing that taxation substantially affects also bank incentives to monitor borrowers.

The reminder of the paper is organized as follows. Section 1.2 describes the dataset and the identification strategy. Section 1.3 presents the results of the empirical analysis. Section 1.4 concludes.

1.2 Data and identification strategy

1.2.1 Data

This paper uses data from the Credit Register (CR) of "Banca d'Italia" (Bank of Italy). This data includes quarterly loan-level information on virtually all business loans extended to limited liability companies in Italy from 2005 to 2016. Since we aim at investigating bank monitoring, we limit our data only to outstanding credit relationships whose length is longer than one quarter. To address our research question, we need to collect information on borrower's characteristics that may affect the likelihood of a repayment. We are able to retrieve such information only for firms included in the CERVED database. This database contains information on balance sheet and income statements of all limited companies operating in Italy. For this reason, we focus our attention on business loans to limited liability companies. Also, for identification purposes, we limit our dataset only to loans granted by small banks operating at the regional level. The reason for that will be clarified in the next subsections.

Our dataset combines loan-level data from the Italian CR with data on firms characteristics, bank conditions, local taxation and macroeconomic factors obtained from different sources. Hereinafter, we describe each group of variables. Table A1.1 in Appendix B provides a detailed summary of this information.

Loan variables

The variables of major interest in our study are those capturing bank monitoring and loan repayment. To build our proxy for bank monitoring we exploit data on requests for information made by banks to the CR on their existing borrowers. Each month banks can submit requests to get information on the total exposure of the banking system to a specific borrower. These requests have different motivations. Our variable for bank monitoring, *Monitor*, is constructed as the total number of requests for information made by a bank on an existing borrower in a given quarter. To build this variable we aggregate all requests for information without making any distinction regarding the reason. Our proxy for bank monitoring and the rationale behind it will be discussed in detail in the next subsection.

We consider four different variables for loan repayment: *Past-due dummy*, a dummy variable equal to one if the loan is past-due by 90 days or more, and zero otherwise; *UTP dummy*, a dummy equal to one if the loan is defined as "unlikely-to-pay", UTP, meaning that the bank envisages the possibility that the loan will not be repaid in full; *Bad loan dummy*, a dummy equal to one if the loan is defined as "bad loan", meaning that the bank considers the loan as impaired; *NPL dummy*, a dummy equal to one if the loan falls into one of the previous categories (nonperforming loan, NPL), and zero otherwise. While the latter variable indicates whether a loan is in general nonperforming, the first three variables capture a condition of distress which is progressively more severe.

Three further loan variables, that we use as controls in our regression analysis, are *Length relation*, *Share guarantee* and *Share exposure*. *Length relation* is the duration of the credit relationship expressed in quarters. *Share guarantee* consists in the value of the credit guarantee (collateral or personal guarantee) as a fraction of the loan. Finally, *Share exposure* is the ratio between the amount of the loan extended by the individual bank and the total borrowing of the firm from the banking system.

All our loan variables are built using information from the CR and are available on a quarterly frequency. In our regressions *Monitor*, *Length relation* , *Share guarantee* and *Share exposure* are included with a lag to make sure that they are predetermined with respect to the dependent variable.

Firm variables

We define a wide set of firms variables capturing relevant firm characteristics that may affect loan repayment. When we do not include firm-time fixed effects in our regression analysis, we use these variables as controls. These include firm size, *Size firm*, profitability, *ROA firm*, the credit score assigned by the provider of the CERVED database, *Credit score firm*, the equity-to-asset ratio, *Capital ratio firm*, and the code identifying the industry in which the firm operates, *Industry firm*. *Size firm* is computed as the logarithm of total assets, whereas *ROA firm* is calculated as the ratio of net income to total assets. *Credit score firm* values range from 1 to 9. A credit score of 1 corresponds to firms with the highest credit quality, while a credit score of 9 corresponds to firms which are essentially in default. *Industry firm* takes 5 different values, each one corresponding to a specific sector. These variables are retrieved from CERVED database and are available on a yearly frequency. In our econometric specifications they are included with a lag to ensure that they are predetermined with respect to the dependent variable.

Bank variables

In our econometric analysis we consider a wide set of bank variables affecting both loan repayment and monitoring incentives. These include bank size, *Size bank*, the equity-to-asset ratio, *Capital ratio bank*, the liquidity ratio, *Liquidity ratio bank*, bank profitability, *ROA bank*, the ratio of nonretail deposits to total deposits, *Nonretail deposit ratio bank*, and the fraction of nonperforming loans to total loans to the private sector, *NPL ratio bank*. *Size bank* is computed as the logarithm of total assets, *Liquidity ratio bank* as the ratio of liquid assets to total assets, and *ROA bank* as the ratio of net income to total assets, respectively.

All these variables are built using information retrieved from the Credit Bureau managed by the Bank of Italy and are available on a yearly frequency. In our regressions bank factors are included with a lag to make sure that they are predetermined with respect to the dependent variable.

Regional variables

One of the most relevant variables in our empirical analysis is the regional IRAP tax rate, *IRAP*. This consists in the IRAP tax rate applied to banks in the bank's region. In our regression analysis we also include variables capturing local macroeconomic conditions of the firm's region that may affect loan repayment. In particular, we consider GDP growth, *GDP growth*, the employment rate, *Employment*, and the inflation rate, *Inflation*. *GDP growth* is computed as the first difference in the logarithm of GDP. Both *Employment* and *Inflation* are expressed in percentage points. In some cases the region of the firm does not coincide with the region of the bank. To control for potential spillover effects from neighboring regions, we also consider similar variables for the region of the bank, that is *GDP growth region bank*, *Employment region bank*, and *Inflation region bank*.

All these variables are available on a yearly frequency. In our econometric specifications they are included with a lag to ensure that they are predetermined with respect to the dependent variable. Data on the local IRAP tax rates come from the Bank of Italy and the Ministry of Economy and Finance. Data on regional macroeconomic conditions are drawn from the Italian National Institute of Statistics, "Istituto Nazionale Statistica" (ISTAT).

1.2.2 Measuring bank monitoring

The most relevant data extracted from the CR concerns the requests for information that banks make on their existing borrowers. Each month, banks can submit these requests in order to get information on the total current credit exposure to a specific borrower by all banks in Italy. In particular, the bank can retrieve information on the amount of each loan granted by other banks to the borrower, as well as on the objective conditions of deterioration of each individual exposure.[3] This information is provided essentially for free, as the cost of one request amounts to few euro cents. In a given month, a bank can submit more than one request for information, each one corresponding to a different reason. The request reason is classified as "historical information", "in-depth information", "credit limit" and "co-signing".

We have to specify that each bank in Italy automatically receives, on a monthly basis, exactly the same qualitative information that can be requested from the CR. There is a difference, though, in terms of quantity of available information that can be obtained. The automatic updated information received from the CR provides a snapshot of the situation at the current month. An actively submitted request to the CR, instead, allows a bank to retrieve also historical information, up to 36 months backward.

It is likely that banks store the automatic flow of information received from the CR in a proprietary database. This raises the question of why banks request information on their existing borrowers in the first place. There are two main motivations for that. First, the bank wants to obtain the most reliable information on current and past records of loans granted to a specific firm by other lenders. This is justified by the fact that data in the CR can be subject to amendments. Indeed, the regulatory guidelines of the CR define in detail how banks should correct erroneous information provided and specify penalties to non-compliers, suggesting that amendments are not uncommon. For this reason, the bank may want to act prudently and make a request to the CR to ensure that it has reliable and updated information on its client. Second, in extraordinary circumstances, the bank may need to verify or rebuild its database containing information on existing borrowers. For example, after a banking M&A the resulting entity may want to check the existing

[3]Objective conditions of deterioration occur, for example, when a loan is overdue. Any discretionary assessment of the bank on the likelihood of repayment are not taken into account.

information on the borrowers of one of the two banks involved, and/or to create a new pool of information.[4]

In the former case, the bank requests information from the CR because it has an interest in assessing the condition of the lending exposure of the banking system to a specific borrower. This, in turn, can be justified by two reasons: the borrower has applied for a new loan, or simply the bank wants to monitor the creditworthiness of the client. This means that only a fraction of requests for information from the CR are actually associated with monitoring purposes. To build our proxy for bank monitoring we use exactly this subset of requests, which is identified thanks to a rigorous cleansing process.

Our original loan-level dataset includes roughly 5.4 million observations on 283,706 credit relationships having a duration greater than one quarter, and involving 225,669 firms and 458 banks. We observe a positive number of requests for information in 11,971 observations, roughly 0.2% of our sample. As a first step, we drop all observations in which a credit relationship is restored after a break (14,507 observations).[5] This ensures that we consider exclusively outstanding loans with a duration greater than one quarter. Second, we discard all observations in which requests for information are driven by exceptional conditions of the bank that have nothing to do with monitoring activity.[6] Finally, we drop all observations in which we observe an increase in credit extended to an existing borrower (790,136 observations). This allows us to eliminate requests for information that are associated with an increase in lending.[7] As we will explain in more detail in Section 1.3, this also ensures that

[4]A third explanation consists in anecdotal evidence suggesting that a request to the CR might be less time consuming than consulting the automatic information received from the CR. Nevertheless, this strictly depends on the internal technical infrastructure of the bank and, hence, we do not consider it as a major motivation.

[5]A break corresponds to a lack of information on a specific bank-firm relationship in the CR for a certain number of quarters. For instance, it could be the case that the firm gets a first loan from the bank. Once the firm pays it off completely, the loan expires and the credit relationship is not reported anymore in the CR. After a certain period of time the firm may apply for a new loan. If this second loan is approved, the bank-firm relationship will show up again in the CR.

[6]To this end, we use a visual inspection aimed at detecting any atypical clustering in requests for information. We identify 243 observations with anomalies in the average number of requests per client made by a bank in a given quarter. These relate to two banks. We drop all the observations pertaining to the pair bank-quarter in which these anomalies occur. Also, we discard all the requests for information made by a bank that has taken part to a M&A during the year of the merger (186 observations).

[7]A request for information not associated with an increase in lending may still be an indication of a

we can properly investigate the causal effect of bank monitoring on loan repayment. In fact, additional credit extended to a firm is likely to influence its ability to meet the repayment schedule in the future, especially in the short run. For this reason, we need to make sure that we discard all observations in which we detect an increase in lending, both in case a request for information is made or not. In this way, we are able to compare the repayment performance of monitored firms versus non-monitored firms, once having controlled for a relevant set of factors.

This stringent cleansing process yields a panel of 4,554,412 observations, corresponding to 280,650 lending relationships over the period 2005-2016. We observe at least one request for information in 4,141 observations, roughly 0.1% of our sample, corresponding to an average of 9 requests per bank. We use the total number of requests for information made by a bank in one quarter to build our novel monitoring variable, that captures in a direct way the effort exerted by a bank in checking the ability of its borrowers to comply with the contractual obligations.

Our proxy is not intended to quantify the whole intensity of bank monitoring, but rather to capture an observable evidence of a broader phenomenon, similarly to the tip of an iceberg. As already pointed out, if the bank wants to verify the condition of total lending to the firm, it can limit itself to consult the automatic flow of information received from the CR. More importantly, the bank can monitor its borrower in different ways, for example by checking the company's financial report, by visiting the firm on site, and by providing advisory services, such as funding management and business planning. It is reasonable to think that these activities are to some extent correlated. For example the bank may use the information on total indebtedness to provide the firm with advises on its financial structure. Therefore, what really matters to us is the dynamics of this variable, which builds on the

rejected application rather than monitoring activity. As consequence, using this subset of requests as a proxy for bank monitoring may lead us to underestimate the effect of bank monitoring on the likelihood of loan repayment. In fact, if these requests for information are exclusively an indication of a loan rejection, we should not find any effect of bank's requests on loan repayment. As we will extensively show in Section 1.3, we actually find a positive effect. Although we cannot exclude that some of the requests for information are due only to a rejected loan application, our findings limit the concern about a possible underestimation. Additionally, in our main specifications we include firm-time fixed effects, which are aimed to capture any observable and unobservable, time varying and time invariant condition of the borrowing firm, including its demand for credit.

idea that the higher the number of requests for information, the stronger is the interest of the bank in assessing the creditworthiness of the borrower, and hence the stronger is the monitoring intensity of the bank. As such, our proxy of bank monitoring resembles the two measures of Gustafson et al. (2017), which consist in the frequency of bank's requests for information on firm's financial reports and field exams of the borrower initiated by the lender. The main difference is that we use, instead, bank's requests for information on other outstanding loans of the firm with the banking system.[8] Our approach to gauge bank monitoring at the loan-level relates also to other measures identified in the literature. These include the frequency with which the bank reviews the internally generated probability of default of the borrower (Plosser and Santos, 2016), the collateral value, the loan spread, the credit rating and the loan limit (Cerqueiro et al., 2016). Our proxy for bank monitoring and those of Gustafson et al. (2017) allow, though, to capture the effort exerted by the bank in monitoring its borrowers in a more direct way. In what follows we show that the dynamics of our variable are consistent with a bank monitoring interpretation.

Appropriateness of our measure

To check the appropriateness of our proxy for bank monitoring we start with a visual inspection. Plot (a) of Figure 1.1 shows that the average number of requests per borrower made by a bank in a given quarter differs across banks, but exhibits a common pattern. Overall, the number of requests for information increases in the first part of the sample up to the recent financial crisis and reaches its peaks in 2008 and 2009. Afterward, it decreases sensibly and stays at a low level until 2012, then it rises again. Looking at higher level of aggregation (plot (b) of Figure 1.1), we see that the average number of requests per client increases by a factor of five from 2005 to 2009, with a sharp acceleration between the third quarter of 2008 and the third quarter of 2009. Consistently with plot (a), the highest value is achieved exactly in the third quarter of 2009. Immediately after, the average number of requests per client decreases, but only for a short period. In fact, as in the case of the Great Recession, requests for information rise again during the second phase of recession

[8] The logic behind our monitoring variable is also similar to that implied by one of the input used by CERVED to construct its measure of firm's credit quality, called "credit score". The credit score by CERVED consists in a rating measure built on the basis of hard and soft information. One of the input used by CERVED is the number of requests for financial reports submitted by banks.

following the sovereign debt crisis in 2012-2013.[9] Overall, this provides a first evidence in favor of our interpretation of requests for information as a proxy for bank monitoring. In fact, it is reasonable to think that banks are more keen on monitoring their borrowers during a period of economic downturn, as borrowers are more likely to miss their repayment schedule. A second important piece of evidence stemming from the figure is that the average number of requests per client exhibits a certain seasonality, with a higher concentration in the first and in the fourth quarter for most years. In particular, the first and the fourth quarter correspond to the highest number of requests for 10 out of 12 years considered in our sample. This is consistent with the idea that banks may want to control the conditions of their borrowers around the balance sheet date, which is the most relevant period of the year for a company.[10]

[Insert Figure 1.1 here]

Monitored firms and monitoring banks

So far we have shown evidence that validates the use of requests for information as a reliable measure of bank monitoring. The next step is to investigate which borrowers are more likely to be monitored and which banks are more likely to monitor. This analysis provides us with further evidence to corroborate the interpretation of our variable as a proxy for bank monitoring.

A first reason why a bank may want to monitor a firm lies in a concern about the ability of the firm to meet its loan obligations. This may occur either before or after a full-blown of payment arrears. Figure 1.2 shows that bank requests for information are related to a nonperforming exposure only in 7.1% of cases. Also, once we move from past-due exposures to higher degrees of distress, namely unlikely-to-pay and bad loans, the percentage of requests for information decreases steadily. Overall, this means that

[9] Italy experienced a negative GDP growth in 2008-2009 and 2012-2013.

[10] Most limited liability companies in Italy set the balance sheet date on December 31 and approve the annual report by the end of the following April. It is reasonable to think that the bank concentrates its monitoring activity towards the balance sheet date and the approval of the annual report as to assess the lending exposure to its clients. Additionally, the bank can retrieve the most meaningful and significant information about the company at this time of the year. In fact, at the balance sheet date the firm has a more clear picture of its revenues and expenditures. As a consequence, it is more likely that the firm takes a decision, either voluntary or forced, to repay its loans in this period of the year.

banks primarily exert monitoring with the intention of preventing firms from missing their repayment schedule. Once a loan is in arrears, the marginal benefit of monitoring decreases with the severity of the distress.

A lack of information about the firm could be a second driver of bank monitoring. Figure 1.2 shows that roughly 10.5% of observations with a positive number of requests is related to credit exposures that are close to the minimum thresholds to be included in the CR.[11] Banks are likely to hold limited information about these loans, as the credit exposure might have become eligible to enter the CR only in recent times. Therefore, this finding suggests that a bank has monitoring incentives if it lacks knowledge about the conditions of the firm.

[Insert Figure 1.2 here]

We now look more closely at the individual features that make a firm more likely to be monitored and a bank more likely to monitor. To this end we perform an econometric exercise which is intended to highlight relevant correlations. In the first specification of Table 1.1 we investigate the role of firm characteristics. Specifically, we regress the number of requests for information made by a bank in the quarter on a set of loan and firm variables capturing the conditions of the borrowing firm. We include macro variables and fixed effects as to control for potential confounding factors. In particular, our set of fixed effects contains bank-quarter fixed effects to control for any observable and unobservable, time varying and time invariant condition of the bank. This ensures that we can focus on the relation between firm factors and bank monitoring.

We find a negative and statistically significant coefficient for *Share exposure*, meaning that the higher the amount of the loan with respect to the firm's total indebtedness, the lower the intensity of bank monitoring. Postulating that the main lender of a firm has access to a greater amount of information, we argue that this variable captures the level of knowledge that the bank has about the firm with respect to other lenders. A more standard measure of bank's knowledge about the firm is *Length relation*, whose coefficient is negative and highly significant as well. This suggests that the intensity of bank monitoring is stronger

[11] The Italian CR requires banks to provide information on credit exposures when specific conditions are met. To define whether an exposure is close to the minimum threshold, we consider the most relevant requirements: (i) the total volume of the credit exposure is greater or equal to €30,000, or; (ii) the credit exposure is defined as bad loan and its volume, net of losses, is greater of equal to €250.

the lower the duration of the lending relationship. We also find a negative and significant coefficient for *Share guarantee*, meaning that the bank has lower monitoring incentives the higher the guarantee as compared to the loan amount. In addition, the positive and statistically significant coefficient of *Credit score firm* shows that firms with a lower credit quality are more likely to be monitored. Interestingly, it seems that banks monitor more large firms. However, the coefficient of *Size firm* is only slightly significant.

The regression in the second column of Table 1.1 improves the preceding by including firm fixed effects to control for any unobservable time invariant condition of the firm. In this way we limit possible concerns of omitted variable bias. The coefficients of *Share guarantee*, *Share exposure*, and *Length relation* are virtually unchanged and, if anything, slightly stronger. The coefficient of *Credit score firm* reverts its sign and looses most of its significance. This is hardly surprising, as this variable captures the creditworthiness of the firm, which is likely to be stable over time. This means that *Credit score firm* might be partially subsumed by firm fixed effects. In contrast to the previous specification, the coefficient of *Size firm* is negative and highly significant. This suggests that the intensity of bank monitoring is stronger the higher the firm's opacity. Interestingly, we find that *ROA firm* is positively correlated with bank monitoring. This somewhat counterintuitive result is fully in line with the theoretical model on bank monitoring presented in Section 1.2.3. Intuitively, once we control for the creditworthiness of the borrower, banks have higher incentives to monitor firms with higher ROA as they can extract higher expected profits from lending. Indeed, firms with high profitability are, unconditionally, more likely to repay. As long as these firms guarantee higher expected profits from lending, banks are more willing to exert a little effort to ensure the repayment of these borrowers rather than devote a great effort to foster the repayment of firms with low profitability. Interestingly, none of the macro variables is statistically significant. Overall, it seems that that, once having controlled for firm characteristics and bank characteristics, macro conditions do not play a relevant role.

In the third specification we extend our analysis investigating whether the length of the credit relationship influences the magnitude of the effect of firm's opacity. To this end, we include the interaction of *Length relation* with *Size firm* among regressors. As before, we find that a longer credit relationship is associated with lower bank monitoring. The coefficient of *Size firm* remains negative, whereas the coefficient of its interaction with *Length relation* is positive and statistically significant. This result reveals that bank monitoring is stronger for firms that are more opaque, but the effect weakens with the duration of the credit relationship. Indeed, banks achieve a deeper knowledge of their borrowers the longer the credit relationship.

We now turn to bank characteristics affecting the incentives to monitor borrowers. In the fourth specification we estimate a model that is symmetrical to those described so far. Specifically, we regress the number of requests for information made by a bank in the quarter on our set of bank variables, including various controls and fixed effects. To make sure that we control for any observable and unobservable, time varying and time invariant characteristic of the firm, we include firm-quarter fixed effects in the specification. This means that we focus on firms having multiple credit relationships and we compare the number of requests for information across banks lending to the same firm.

Share exposure, *Length relation*, *Size bank*, and *Nonretail deposit ratio bank* are the only variables that turn out to be statistically significant. The negative and significant coefficients of *Share exposure* and *Length relation* confirm that banks having a better knowledge of their borrowers are less likely to monitor. Also, the positive and significant coefficient of *Size bank* suggests that large banks are less likely to monitor. As for *Nonretail deposit ratio bank*, although this variable was intended to estimate the effect of unsecured deposits, the negative and significant coefficient is likely to capture a size effect.[12]

Most of the coefficients of the other factors are in line with expectations, but are not significant. For example, the positive coefficient of *ROA bank* and the negative coefficient of *NPL ratio bank* are consistent with the theoretical predictions presented in Section 1.2.3. A high profitability implies a low bank's probability of default, whilst the opposite applies to the ratio of nonperforming loans. A low default probability, in turn, entails high expected profits to shareholders stemming from lending. Thus, our model suggests that bank stability improves bank incentives to monitor borrowers. The sign of *ROA bank* and *NPL ratio bank* are exactly in line with this intuition. Additionally, the negative coefficient of *Liquidity ratio bank* is in line with the idea that banks holding a high amount of liquid assets are able to

[12] If a high value of *Nonretail deposit ratio bank* implies a low fraction of secured deposits, we would expect that the higher *Nonretail deposit ratio bank* the higher bank incentives to monitor borrowers. The reason behind lies in the market discipline exerted by unsecured depositors, as suggested by Diamond and Rajan (2000). However, this results seems to suggest that *Nonretail deposit ratio bank* rather captures the size of the bank. In fact, a bigger bank is likely to have a higher fraction of nonretail deposits.

take on more risk, as they can easily absorb potential losses. Finally, the negative coefficient of *GDP growth region bank* is consistent with the evidence of Figure 1.2, namely banks have higher monitoring incentives during periods of economic downturn. Nevertheless, as already pointed out, these coefficients are not statistically different from zero. This findings, as well as the results of the previous specifications, suggest that firm factors play a prominent role than bank factors in driving bank monitoring.

This econometric exercise allowed us to identify in a straightforward way firm and bank characteristics that are correlated with the intensity of bank monitoring. All our results are consistent with a monitoring interpretation of our novel variable based on requests for information from the CR. We conclude that, overall, our findings provide support to our methodological approach in measuring bank monitoring.

[Insert Table 1.1 here]

1.2.3 Identification strategy

Which is the effect of bank monitoring on loan repayment? Assessing this causal relation is challenging. The repayment performance of a firm is likely to influence bank monitoring as well, exposing to the threat of reverse causality. Also, unobservable conditions of the borrowing firm can potentially affect its ability to meet the contractual obligations, making it difficult to identify the effect of bank monitoring in a precise way.

To address our research question we rely on a robust identification strategy that builds on two main pillars: first, we exploit taxation as a source of exogenous variation in bank monitoring; second, we use firm-time fixed effects to control for time varying and time invariant, observable and unobservable conditions of the firm.

Taxation is likely to affect bank incentives to monitor borrowers through different channels. We focus on the corporate tax rate and we develop a simple theoretical model to highlight the different mechanisms at play. Our model indicates that an increase in the corporate tax rate entails a decrease in bank monitoring, and vice-versa. Then, relying on this prediction, we focus on small banks and we use regional variation in the IRAP tax rate to retrieve exogenous variation in bank monitoring.[13] In this way, we are able to investigate

[13] We borrow this identification strategy from Bond et al. (2016) and Gambacorta et al. (2017), who exploit small banks and regional variation in the IRAP tax rate to analyze the effects of taxation on bank capital structure.

the causal effect of bank monitoring on loan repayment.

As for possible confounding factors, we focus our analysis on business loans to limited companies exactly because we have extensive information on these firms that we can use to control for the borrower's conditions.[14] Nevertheless, even controlling for observable firm's characteristics, there could still be unobservable factors driving the repayment performance of a loan. For this reason, we saturate our main regressions including firm-time fixed effects. This is made possible by the fact that almost 13% of firms in our sample have multiple credit relationships. As a result, we estimate the causal effect of bank monitoring on loan repayment by comparing the repayment performance of different loans granted by different banks to the same firm at a certain point in time. This means that identification stems from different tax rates applied to banks operating in different regions and lending to the same firm.

Hereinafter, we discuss our first pillar more in detail and we describe the empirical methodology adopted.

A model of taxation and bank monitoring

We develop a simple model of bank monitoring, extending the one of Dell'Ariccia et al. (2014) by introducing a corporate income tax applied to bank profits. Specifically, we consider a representative bank funded only by equity, with fraction k, and deposits, with fraction $1-k$, which operates in a perfectly competitive environment. The bank uses its sources of financing exclusively to grant an arbitrary amount of indistinguishable loans, $L(r_L)$, where r_L denotes the lending rate. The bank faces a downward sloping demand curve, $L'(r_L) \leq 0$. A corporate income tax is applied on revenues from lending, with τ being the tax rate.

Since loans are risky, the bank needs to monitor its borrowers in order to prevent a potential default. The bank possesses a monitoring technology that allows to exert a monitoring effort q, which also represents the probability of loan repayment. Clearly, monitoring does not come for free and entails a certain cost for the bank, $\frac{1}{2}cq^2$, per unit of lending.[15]

[14]This information would be missing for households.

[15]In our empirical setup we use bank requests for information from the CR as a proxy for bank monitoring. We have highlighted that each request costs only few euro cents. Nevertheless, this does not mean that monitoring is costless. In fact, bank monitoring involves a wide spectrum of activities that go beyond the assessment of the information owned by the CR. These include checking the firm's financial report,

There is no deposit insurance, and both shareholders and depositors are assumed to be risk-neutral. As such, they require a return that compensates their opportunity cost. The rate of return crucially depends on the probability of loan repayment and equals $r_E = \frac{r+\xi}{q}$ for shareholders and $r_D = \frac{r}{\mathbb{E}[q|k]}$ for depositors, with r being the risk-free interest rate and ξ a positive equity premium.

We further introduce a friction affecting bank capital structure. We assume that the interests paid on deposits are tax deductible, in line with Gambacorta et al. (2017). This distortion implies that equity is a less convenient source of funding than deposits.

The bank determines the optimal lending rate, r_L^*, the optimal capital structure, k^*, and the optimal monitoring effort, q^*, as to maximize the expected profits:[16]

$$\underbrace{\max}_{r_L,k,q,0<q\leq 1} \Pi = \left\{ q\left[\left(r_L - r_D\left(1-k\right)\right)\left(1-\tau\right) - r_E k \right] - \frac{1}{2}cq^2 \right\} L\left(r_L\right) \tag{1.1}$$

Note that in the maximand the cost of bank monitoring does not reduce taxable income. This is consistent with a view of bank monitoring as a non-pecuniary effort exerted by loan officers in assessing and improving the likelihood of loan repayment. However, it is reasonable to think that bank monitoring involves also monetary costs, for example in terms of remuneration of loan officers. As we will discuss in the next section, our empirical setup exploits the IRAP tax applied to banks, whose tax base includes both profits and wages. This implies that the pecuniary costs supported by Italian banks for monitoring purposes do not reduce, but rather increase the IRAP taxable income. Thus, the way we model the costs of bank monitoring is consistent also with the framework of our empirical analysis.

Solving the model provides us with relevant insights. An increase in the corporate tax rate entails three main effects: (i) net profits decrease because of higher tax expenditures; (ii) the capital ratio drops as equity funding becomes less attractive; (iii) the lending rate increases as a result of a shift of tax burden from the bank to its borrowers. The first two effects entail a decrease in bank monitoring, which is only partially counteracted by the latter effect. Hence, overall an increase in the corporate tax rate leads to a decrease in bank monitoring, as stated in Proposition 1.

performing field exams, visiting the firm on site, providing advisory services etc. All these activities require substantial pecuniary and non-pecuniary costs for the bank.

[16] There is no agency conflict between bank managers and shareholders as their interests are assumed to be perfectly aligned.

Proposition 1. Equilibrium bank monitoring decreases with the corporate tax rate, $\frac{\partial q^*}{\partial \tau}$.

Indeed, the resulting optimal level of monitoring effort and its derivative with respect to the corporate tax rate are:

$$q^* = \sqrt{\frac{2r\,(r+\xi)^2\,(1-\tau)}{c\,(3r\xi + r^2 + 2\xi^2 + r^2\tau + r\tau\xi)}} \tag{1.2}$$

$$\frac{\partial q^*}{\partial \tau} = -2\,(r+2\xi)\,(r+\xi)\sqrt{\frac{r^3}{c\,(1-\tau)\,(3r\xi + r^2 + 2\xi^2 + r^2\tau + r\tau\xi)^3}} < 0 \tag{1.3}$$

The proof is provided in Appendix A. This result is in line with a classical "skin in the game" argument, in what is suggests that lower expected rents from lending reduce bank incentives to ensure loan repayment by monitoring its borrowers. Since an increase in the tax rate worsens bank stability, our result is consistent with the literature that points to a negative relation between bank stability and risk-taking (Allen et al., 2011; Mehran and Thakor, 2011; Dell'Ariccia et al., 2014).

Further extensions of our model suggest that, when the capital structure is exogenous, the effect of taxation on bank monitoring is stronger for lowly capitalized banks. Moreover, if deposits are fully insured, bank monitoring incentives are completely insensitive to the corporate tax rate.[17]

IRAP

IRAP is a flat tax on the value added generated by firms and public administrations that was introduced in Italy in 1998.[18] Until 2001 the IRAP tax rate was the same across Italian regions. Since 2002 each region is allowed to set its local IRAP tax rate, increasing or decreasing the national basic rate by maximum one percentage point until 2008 and 0.92 percentage points since then. The IRAP tax rate applied to banks usually differs from that applied to other firms. Typically, the former is larger and has been subject to a higher variation over time than the latter. Revenues from the IRAP tax are mainly used to finance

[17] These results are available upon request.

[18] The difference between the IRAP tax and a standard corporate income tax lies in the tax base. For example, for the specific case of the IRAP tax applied to banks, the tax base includes not only profits but also wages.

the National Health Service ("Servizio Sanitario Nazionale", SSN),[19] which is organized under the Ministry of Health and administered at the regional level. For instance, in 2012 revenues from the IRAP tax represented about 30% of the total funding of the National Health Service (MEF, 2012).[20] While in normal times regions are free to modify the local IRAP tax rate within the range limit, if a health care deficit occurs the regional IRAP tax rate is automatically increased ex lege. In our sample period, this happened five times, specifically in Abruzzo in 2006, Campania in 2006 and 2010, Calabria in 2010, Lazio in 2010, and Molise in 2010.

Table 1.2 reports the regional IRAP tax rates applied to banks during our sample period.[21] We detect 59 changes in the local IRAP tax rates occurred between 2006 and 2016, 35 increases and 24 decreases. This guarantees that we are able to exploit a significant variation in the IRAP tax rate both across regions and over time. Since revenues from the IRAP tax are mainly used to finance national health care expenditures, the IRAP tax rates are reasonably orthogonal to firm's conditions affecting the likelihood that the loan is repaid, as well as to bank monitoring incentives. This guarantees the exogenity of the IRAP tax rates, meaning that our identification strategy is appropriate and sound. One concern, though, could be that local macroeconomic conditions jointly affect firm's loan repayment, bank monitoring and local IRAP tax rates. For example, during an economic downturn firms might be unable to meet their credit obligations, requiring banks to monitor more intensively; at the same time, local governments may increase the IRAP tax rates in response to a reduction in other sources of funding of the National Health Service. To limit this concern, we always control for relevant macroeconomic conditions in our estimation models. More importantly, in Section 1.3.2 we conduct an exercise to verify whether local IRAP tax rates depend (at least linearly) on regional macro conditions and aggregate bank factors. We find no evidence of such dependence, corroborating our identification strategy. A further concern could be that there is a correlation between the IRAP tax rate applied

[19] Article 38 of the Legislative Decree No. 446 of 15 December 1997 states that 90% of revenues from the IRAP tax, net of the quota allocated to the State, are used to finance national health care expenditures.

[20] This corresponds to roughly €38 billions. Such substantial amount is due to the fact that the Italian National Health Service, which provides healthcare to all citizens and residents in Italy, is funded totally by tax revenues.

[21] Despite our dataset covers the time period 2007-2016, we report the tax rates also for 2006, as in our econometric analysis we use the lagged value of the IRAP tax rate.

to banks and the IRAP tax rate applied to firms, with the latter being able to affect significantly the firm's ability to repay. In our sample we observe that only 35 of changes in the IRAP tax applied to banks are concomitant with changes in the IRAP tax rate applied to limited companies, meaning that a substantial fraction of the variation concerns only banks and is not affecting firms. More importantly, our methodology focuses on variation within firm-time, as we include firm-time fixed effects in our econometric specifications. This feature is crucial to control for any relevant condition of the firm affecting its likelihood to repay, including the tax burden.

[Insert Table 1.2 here]

Local banks

Banks that operate in different regions determine the IRAP tax base as a weighted average of the local tax bases calculated in proportion to the amount of deposits held in each region. For this reason, we cannot exploit the local IRAP tax rates as a source of exogenous variation in the monitoring intensity of big banks operating on a national scale. However, this is possible for small banks that operate at the local level. In fact, local banks are typically subject to special regulatory restrictions, implying that they cannot belong to a banking group and must operate in a very limited geographic area. As such, these banks are mostly active in one region. This means that changes in the IRAP tax rate affect the whole tax base, which is why we focus the attention on such local banks. Moreover, up to 2011 local banks were almost exclusively subject to the IRAP tax. As such, the IRAP tax rate exerts a relevant influence on their behavior. Unlike non-financial firms, banks can deduct interest expenses from the IRAP tax base. This implies that changes in the IRAP tax rate have an impact on the capital structure of local banks, as documented by Bond et al. (2016) and Gambacorta et al. (2017). In our context this is likely to play a role in affecting bank monitoring incentives, as suggested by our theoretical model.

To assign each bank to one region we look at the region in which the bank has most of its branches.[22] This approach is sound, as 99% of local banks in our sample has a number

[22] Bond et al. (2016) and Gambacorta et al. (2017) look, instead, at where the bank is headquartered. Although the outcome is likely to be almost identical, we consider our approach more reliable as the IRAP tax base is determined in proportion to the amount of deposits held in each region.

of branches in the first region of major activity at least 1.5 times as large as the number of branches in the second region.

Although local banks are subject to a special regulation, which influences the composition of their balance sheet,[23] they experience similar levels of profitability to those of big banks (Bond et al., 2016; Gambacorta et al., 2017). In light of that, there is no reason to believe that local banks' monitoring incentives respond in a different way from those of big banks to changed economic conditions.

Empirical methodology

Our empirical strategy includes two main estimation models. The first consists in a 2SLS regression in which we use the IRAP tax rate as instrument for our proxy for bank monitoring. This exercise allows to assess which is the impact of an increase in the requests for information, driven by a change in taxation, on the repayment performance of a loan.

By construction, our proxy for bank monitoring may underestimates the total monitoring effort exerted by a bank. The theoretical model presented in Section 1.2.3 suggests that the corporate tax rate affects bank incentives to monitor borrowers on the whole. Hence, to circumvent this issue, we exploit variation in the IRAP tax rates as a way to capture changes in the overall monitoring activity of banks. Specifically, we estimate a second reduced form model in which we quantify how an increase in the overall bank monitoring driven by taxation affects loan repayment.

1.2.4 Regression dataset

Descriptive statistics

To build the final dataset used in our regression analysis, we merge the loan-level data from the Italian CR with data on firms characteristics, bank conditions, local taxation and macroeconomic factors retrieved from different sources. Due to a lack of availability for some variables over specific time periods, we lose a substantial amount of observations.[24] Additionally, we discard not only credit relationships having a duration of one quarter, but also those having a duration of two and three quarters. This is necessary because, in most

[23] For instance, at least 50% of assets of these banks has to be represented either by risk-free assets or loans to shareholders. Also, a high fraction of their profits has to be retained in reserves.

[24] For example, data on bank balance sheet are available starting only in 2006.

specifications, we include our measure of bank monitoring lagged of two quarters. Hence, we need to ensure that we take into account only the requests for information made with respect to existing borrowers having an established credit relationship with the bank. For the same reason, we also drop all observations in which we observe a break in the credit relationship in the current quarter or in the previous two. It results a panel of 2,445,744 observations that encompasses 217,199 credit relationships involving 177,781 firms and 446 banks. In this panel credit relationships are the cross-sectional unit and years 2007-2016 are the time unit.[25] In the main regressions of our econometric analysis we rely on firm-time fixed effects to control for any observable and unobservable, time varying and time invariant condition of the firm. To this end, we exploit the fact that roughly 13% of firms in our sample have multiple credit relationships. The reduced sample including only firms that have multiple credit relationships is a panel of 556,717 observations. This encompasses 53,776 credit relationships involving 23,393 firms and 440 banks over the time period 2007-2016.[26] Table 1.3 reports summary statistics of the variables used in our regression analysis for both the full sample and the reduced sample.

Variation exploited

The main results of our econometric study are obtained using the reduced sample pertaining to firms having multiple credit relationships. The inclusion of firm-time fixed effects in our specifications is an essential ingredient of our methodology. Investigating the causal effect of bank monitoring on loan repayment requires to control of any condition of the firm that may influence the likelihood that the firm repays its loan. In our regression analysis we consider a wide set of firm variables. Nevertheless, even the use of an arbitrarily large set of controls does not prevent that we miss some relevant unobservable factors. This is why we believe that firm-time fixed effects are essential. As noted above, when we include this set of fixed effects we work on a dataset that is roughly one fourth of the original full sample, as we are focusing on firms having multiple credit relationships. More importantly, our identification strategy based on exogenous changes in the IRAP tax rates builds on a specific kind of variation that we need to discuss in detail. In particular, our identification comes chiefly

[25]The full sample corresponds to the sample used in regressions (1)-(4) of Table 1.4.

[26]The reduced sample corresponds to the sample used in the regressions of Table 1.6. In the following tables there is a deviation from the number of 556,717 observations depending on the set of variable included in the econometric specification.

from firms having multiple credit relationships with banks established in different regions and, hence, subject to different tax rates. To draw any implication for the external validity of our results we need to understand how representative these banks and these firms are.

Figure 1.3 shows the distribution of firms and banks across regions in the full sample and the reduced sample. Looking at the full sample, we note that firms and banks are distributed throughout the whole country, but the majority is located in the center-north regions. This remains true even in the sample including only firms having multiple credit relationships. More importantly, despite the number of firms is reduced by a factor of 7 in the reduced sample, the distribution of firms and banks across regions is virtually unchanged. This is important evidence suggests that the reduced sample is sufficiently representative. This claim is corroborated also from the statistics reported in Table 1.3. In fact, all variables show similar values in both samples.

The next step is to assess precisely which is the main source of variation that we exploit. First, we observe that 14% of firms included in the reduced sample borrow, at least in one quarter, from one or more banks located in a different region from their own. We also detect that 29% of banks in the sample lend, at least in one quarter, to firms located in a different region from their own. Figure 1.3 shows that the distribution of these firms and these banks is similar to what observed for all firms and banks included in the reduced sample. A natural question arises: is it that firms move to borrow from banks located in a different region, or rather is it that banks establish branches outside their region and lend to firms located in other regions? We cannot answer precisely to this question as we do not have information about the branch of the bank where the credit relationship takes place. However, we can provide some useful numbers. Specifically, we find that 62% of observations in which the region of the firm differs from that of the bank are associated to banks that operate only in one region. In these cases, it is certainly the firm which moves in order to borrow from a bank located in a different region. Thus, we conclude that most of the variation is driven by firms that seek credit even outside their region. It is likely that these firms operate close to the border. For this reason, in our specifications we control for macroeconomic factors of both the firm's region and the bank's region.

[Insert Figure 1.3 here]

As a last step, we check if firms included in our sample are representative of the whole universe of firms in Italy. Looking at the number of employees and the size of the asset side, we observe that most of these firms fall under the definition of small and medium-sized enterprises (SMEs). This is not surprising, as we focus on firms borrowing from local banks operating at the regional level. Moreover, according to the SMEs CERVED Report 2014, SMEs represent about 96% of the total number of firms operating in Italy. To assess if our sample of firms is sufficiently representative, we compare the statistics of Table 1.3 with those of the SMEs CERVED Report of 2014. Clearly, our sample covers a period of 10 years characterized by varying macroeconomic conditions, while the Report refers only to the year 2014. However, this comparison is still useful to understand if the order of magnitude of our variables is consistent with what observed on the national scale. We find that the mean value of the credit score is very close to what described in the SMEs CERVED Report of 2014, but firms in our sample exhibit lower profitability and a lower level of capitalization.

[Insert Table 1.3 here]

1.3 Results

1.3.1 Preliminary analysis on bank monitoring and loan repayment

We begin with a graphical inspection of the relation between bank monitoring and loan repayment. Figure 1.4 shows the percentage of nonperforming loans in each quarter against the average number of requests for information submitted by banks two quarters before. We focus on the relation between requests for information and loan distress two quarters ahead because, as we will show more in detail in Section 1.3.2, this is the horizon where we identify the stronger effect of bank monitoring on loan repayment. Despite the high dispersion, this plot documents that bank requests for information are negativity related to nonperforming loans, suggesting that bank monitoring may have a positive effect on loan repayment

[Insert Figure 1.4 here]

This finding is confirmed once we analyze more granular data at the loan level. The first four columns of Table 1.4 report the results of an exercise in which we regress a dummy variable for different types of nonperforming loans on our variable for bank monitoring lagged of two quarters. In the first specification, we include a wide set of controls spanning loan, firm, bank and macroeconomic factors. Specifically, *Share guarantee*, *Share exposure*, *Length relation*, *Size firm*, *ROA firm* and *Credit score firm* capture loan and firm characteristics that may influence the ability and the willingness of the firm to repay its loan. *GDP growth region firm*, *Employment region firm*, *Inflation region firm*, *GDP growth region bank*, *Employment region bank*, and *Inflation region bank* control for macroeconomic conditions of the firm's region and the bank's region that may affect the business environment of the firm and, hence, its repayment performance. Finally, *Capital ratio bank*, *ROA bank*, *NPL ratio bank*, *Size bank*, *Liquidity ratio bank* and *Nonretail deposit ratio bank* capture bank conditions that may influence bank incentives to recognize the loan as nonperforming.[27] All regressors are lagged according to their frequency, so as to ensure that each control is predetermined with respect to the dependent variable and, at most, concomitant with our proxy for bank monitoring.

The coefficient of *Monitor* suggests that one additional request for information is associated with a decrease of 0.8 percentage points in the likelihood that the loan becomes nonperforming two quarters ahead. Since we rely on a severe multiclustering at the year-quarter and bank level, the statistical significance of the effect is substantial. When we disaggregate the dependent variable into different dummies capturing an increasing degree of loan distress, we find a negative relation for the last two categories, but the correlation is somewhat stronger for bad loan exposures. Overall, these results are in line with what observed in Figure 1.4.

If we focus on the other covariates of the first model we see that loan distress is positively associated with the ratio of loan amount to total firm's borrowing and with the length of the credit relationship. As expected, firms with low profitability and bad credit quality, as well as small firms, are more likely to miss the repayment schedule. Also, good macroeconomic conditions in terms of employment rate are correlated with a lower probability that the loan becomes nonperforming. More controversial is the interpretation of the other variables whose coefficient is statistically significant.[28]

[27] The recognition of a credit exposure as past-due is rather mechanical and occurs each time a loan is past-due by 90 days or more. Banks have, instead, greater flexibility in classifying a credit exposure as unlikely-to-pay or bad loan, as this depends on a subjective assessment.

[28] For example, the coefficients of *Capital ratio bank* and *ROA bank* suggest that a nonperforming loan is positively related with a low level of capitalization and low profitability of the bank. This reveals a certain

As a next step, we investigate what could be the channels through which bank monitoring is associated with improved loan repayment. Models (5)-(6) analyze the relation between *Monitor* and *ROA firm*, *Capital ratio firm* and *Credit score firm*, respectively. First, we find that bank monitoring is positively related to future firm's capitalization. This suggests that bank monitoring is not limited to a mere costly state verification (Townsend, 1979), but encompasses also other activities that may improve the conditions of the firm. We also find a positive coefficient for *ROA firm* and a negative coefficient for the credit score, suggesting a positive correlation also with firm's profitability and credit quality, but the coefficients are not significant. Overall, these findings reveal that bank monitoring may entail a series of actions that improve the conditions on the firm and, consequently, its repayment performance. We leave this analysis for future research.

The last model extends the first regression replacing firm variables with firm-time fixed effects to control for any observable and unobservable condition of the borrowing firm that may affect its ability to meet the repayment schedule. The coefficient of *Monitor* reverts sign.[29] Despite the extensive amount of controls and the fact that we include our measure of bank monitoring lagged of one quarter to limit reverse causality, we should not be tempted to interpret these results from a causal perspective. In fact, in these regressions there is still room for potential endogeneity. If banks monitor more intensely loans that are more likely to become overdue, then the coefficient of *Monitor* can be biased upward. In the next subsection we discuss how we deal with this issue and we present the main results of our empirical analysis.

[Insert Table 1.4 here]

1.3.2 Main results

To identify the causal relation between bank monitoring and loan distress we rely on our identification strategy, exploiting exogenous variation in the regional IRAP tax rates. We first show some necessary exogeneity checks to corroborate our methodological approach. Afterward, we present the main regression analysis.

degree of reverse causality.

[29]For the sake of brevity, we do not report additional specifications for each type of distress. The results, though, confirm what observed in regression (8). Concerning specifications (5)-(7), we cannot run similar regressions including firm-time fixed effects, as the dependent variable is invariant within the fixed effect.

Exogeneity checks

As a preliminary check, we verify if the IRAP tax rate is indeed exogenous to bank monitoring and loan repayment. Table 1.5 reports the results of different specifications in which we regress the IRAP tax rate on a set of local macroeconomic conditions and bank variables. The former encompasses the same variables used insofar for the bank's region, whereas the latter include the aggregate capital ratio, *Capital ratio region bank*, and ROA, *ROA region bank*, of the banking system at the regional level, as well as the average ratio of nonperforming loans of banks operating in a specific region, *NPL ratio region bank*. We also include among the independent variables the basic IRAP tax rate defined at the national level and a dummy variable equal to one if an increase in the IRAP tax rate occurs in response to a regional health deficit, Δ*IRAP health*. We find that the IRAP tax rate depends exclusively on the current basic IRAP tax rate at the national level and on the event of a regional health deficit. Neither macro factors nor aggregate conditions of the banking system at the regional level correlate with the IRAP tax rate. Although these results cannot rule out other kind dependence than the linear one, they provide evidence that supports our identification strategy. As a last remark, we point out that the national basic tax rate is likely to depend on aggregate macroeconomic conditions of Italy. In our main regression models we always include firm-time fixed effects, which means that we actually control for the situation of whole Italian economy. In other words, our identification crucially depends on variation in the IRAP tax rate across regions, which is exogenous as it is driven by differences in healthcare expenditures.

[Insert Table 1.5 here]

Bank monitoring and loan repayment: A 2SLS approach

We can now describe the baseline model adopted to investigate the causal link between bank monitoring and loan repayment. Our methodology relies on instrumental variables and consists in estimating the following 2SLS model

1st Stage:

$$\text{Monitor}_{i,b,r,t-2} = \alpha + \beta \text{IRAP}_{r,t-6} + \boldsymbol{\gamma' X}_{i,b,r,t-n} + \mu_{i,t} + \mu_b + \mu_r + \varepsilon_{i,b,r,t-2} \tag{1.4}$$

2nd Stage:

$$\text{NPL dummy}_{i,b,r,t} = \alpha + \beta \widehat{\text{Monitor}}_{i,b,r,t-2} + \boldsymbol{\gamma'X}_{i,b,r,t-n} + \mu_{i,t} + \mu_b + \mu_r + \varepsilon_{i,b,r,t} \tag{1.5}$$

where $\text{Monitor}_{i,b,r,t-2}$ is the number of requests for information made by bank b operating in region r on firm i at time $t-2$; NPL $\text{dummy}_{i,b,r,t}$ denotes a dummy variable equal to one if the credit exposure of bank b to firm i is in distress at time t, and zero otherwise; $\text{IRAP}_{r,t-6}$ is the tax rate applied to the bank, that we use as instrument for bank monitoring (so called "excluded instrument"). $\boldsymbol{X}$ stands for a vector of controls, encompassing loan characteristics (*Share guarantee*, *Share exposure* and *Length relation*), macro regional conditions (*GDP growth region bank*, *Employment region bank* and *Inflation region bank*), and bank variables (*Capital ratio bank*, *ROA bank*, *NPL ratio bank Size bank*, *Liquidity ratio bank* and *Nonretail deposit ratio bank*) affecting bank incentives to monitor borrowers as well as the likelihood of loan distress. $\mu_{i,t}$, μ_b and μ_r denote firm-quarter, bank and bank's region fixed effects, respectively. $\widehat{\text{Monitor}}_{i,b,r,t-2}$ is the linear projection of $\text{Monitor}_{i,b,r,t-2}$ onto all the exogenous variables, namely the excluded instrument and controls. Also, both *IRAP* and the control variables are lagged to ensure that they are predetermined with respect to the dependent variable in each stage. $t-n$ represents the lag of the regressor expressed in quarters and depends on its frequency. This model allows us to estimate the causal effect of an increase in bank monitoring, as captured by our proxy, on the likelihood of loan repayment two quarters ahead. We consider a lag of two quarters because this represents the horizon in which we identify the strongest effect, as evinced in Table 1.7.[30]

This 2SLS model is perfectly identified, as we have a single instrument, *IRAP*, for a unique endogenous variable, *Monitor*. To ensure that the IV estimator is unbiased we need to assess if our instrumental variable satisfies the necessary requirements in terms of relevance and exclusion restriction. Economic considerations (i.e. revenues from IRAP are mainly used to finance healthcare expenditures at the regional level), as well as the evidence that the IRAP tax rate is uncorrelated with local macroeconomic conditions and aggregate bank factors at the regional level (Table 1.5) suggest that *IRAP* is undeniably exogenous to bank monitoring and loan repayment. As for the relevance requirement, we rely on standard econometric tests to assess if our instrument is strong enough.

One additional thing to highlight about our model is that, since we include firm-time

[30] In Table 1.7 we report the results of our 2SLS model for different horizons.

fixed effects, identification is mainly provided by loans to the same firm granted by two or more banks operating in different regions and, hence, subject to different IRAP tax rates. Also, this 2SLS model is run on a sample of loans having a length greater than three quarters, in which we have dropped all observations associated with an increase in lending to an existing borrower in the current quarter and/or in the previous two. This ensures that we are investigating in a proper way the effect of bank monitoring on loan repayment. In fact, suppose that in quarter $t-2$ the bank makes a request for information that is associated with new credit granted to an existing borrower. It is likely that the new loan affects the firm's ability to meet its repayment schedule at quarter t. Our cleansing process allows us to get rid of such observations. In this way, we are able to compare the repayment performance of monitored firms versus non-monitored firms, without worrying about the effect of an increase in lending.

Table 1.6 reports the result of this 2SLS regression analysis. The first stage (column 1) highlights that the coefficient of the IRAP tax rate is negative and statistically significant. Consistently with Proposition 1 of our model presented in Section 1.2.3, this finding reveals that an increase in the IRAP tax rate implies a decrease in bank monitoring. This result is particularly striking if one thinks that the tax rate has only a second order effect on bank monitoring, as highlighted in our model. The magnitude of the effect, though, is rather limited in absolute terms. One percentage point decrease in the IRAP tax rate leads to an increase of 0.004 in the number of requests for information made by the bank. But this is not surprising as the average number of requests for information detected in our sample is 0.001 and the standard deviation is 0.029 (panel B of Table 1.3).[31] Thus, the negative effect of the tax rate on bank monitoring is actually substantial in relative terms. In fact, a one percentage point decrease in the IRAP tax rate implies an increase in the number of requests for information that corresponds to four times its average in the sample.

To assess whether *IRAP* fulfills the relevance requirement we perform some standard underidentification and weak identification tests. The Kleibergen-Paap rk LM statistic (4.00) leads us to reject the null hypothesis that the model is underidentified at 95% level. This means that our instrument, *IRAP*, is sufficiently correlated with the endogenous regressor, *Monitor*. Also, the Cragg-Donald Wald F-statistic (18.11) and the Stock and Yogo (2005)

[31] This is exactly why we argue in Section 1.2.2 that our variable of bank monitoring captures only a limited fraction of the overall monitoring activity conducted by the bank.

critical value for a 5% level test that the maximum size of the Wald test is no more than 10% (16.38) suggest that our instrumental variable is strong enough. Nevertheless, both the Cragg-Donald Wald F-statistic and the critical value are meaningful under the assumption of i.i.d. errors. This condition is likely not to hold, which is the reason why we cluster standard errors to draw reliable inference. Thus, we better focus on the Kleibergen-Paap Wald F-statistic, which is cluster-robust. Despite we do not have a specific critical value for this statistic, its value (4.94) is relatively low, especially if confronted with the rule of thumb of 10 suggested by Staiger and Stock (1997). In light of that, we should be careful in concluding that our instrument satisfies, to an acceptable degree, the relevance requirement. For this reason, as suggested by Andrews et al. (2018), we report the results of the Anderson-Rubin test, which allows to derive weak-identification-robust inference.

As for the other covariates, we get similar results to those of Table 1.1. Specifically, bank monitoring decreases with the duration of the credit relationship. This is in line with the idea that banks monitor less the firms that they know better. Also, higher credit guarantee with respect to the loan amount entails a lower intensity of bank monitoring. Finally, a higher employment rate in the bank's region is associated with higher bank monitoring.

Looking at the results of the second stage regressions (columns (2)-(7)), we detect a strong negative effect of bank monitoring on the likelihood that the loan becomes nonperforming two quarters ahead. We interpret the results keeping in mind that our estimates represent essentially a weighted average of local average treatments effects (Heckman et al., 2006; Angrist and Pischke, 2009; Cornelissen et al., 2016).[32] Specifically, we find that an increase in the number of requests for information, that corresponds to a 1 percentage point decrease in the IRAP tax rate, entails a decrease by 3.9 percentage points in the probability that the loan ends up in distress two quarters ahead. This effect is heavily statistically and economically significant, especially in light of the fact that the probability of a loan being nonperforming is 11% in our reduced sample (Table 1.3). Moreover, an increase in the IRAP tax rate of 1% is a rather realistic event. In fact, in our sample period, we observe five times an absolute change in the regional IRAP tax rate that is greater or equal to 1 percentage point (Table 1.2). As mentioned earlier, we perform the Anderson-Rubin test

[32]Our endogenous variable is a count variable. Thus, it can be considered a treatment effect with different levels of treatment (Angrist and Pischke, 2009). As a consequence, we need to interpret our result taking into account a reasonable change in the instrumental variable.

to derive weak-identification-robust inference. This test is similar to a standard t-test, as it allows to assess if the coefficient of the endogenous regressor is statistically different from zero, but is robust to the use of a weak instrument. The Anderson-Rubin Wald statistic confirms that the effect of bank monitoring on loan repayment is statistically significant and the extent of the significance is even higher than that of a standard t-test (5% versus 10%). Looking at the control variables, we see that a lower *Share exposure* and better macroeconomic conditions in terms of GDP are associated with a higher probability of loan repayment.

When considering increasing levels of loan distress, we note that the magnitude of the effect strengthens from past-due to unlikely-to-pay exposures, whereas it actually reverts for bad loans. This finding, as well as the evidence that the coefficient of *Monitor* is significant only in the fourth specification, suggest that the positive impact of bank monitoring on loan repayment is mainly driven by the ability of the bank to prevent a loan from becoming unlikely-to-pay. Additionally, once a loan is in a hopeless condition of distress, bank monitoring seems not to be helpful anymore to foster loan repayment. There could be two different explanations for the latter result. First, when a credit exposure is severely distressed the bank does not have incentives to monitor anymore. Alternatively, the bank still monitors the firm but monitoring is not effective. Looking at Figure 1.2, we observed that the number of requests for information decreases substantially from past-due exposures to bad loans. This evidence provides support to our first conjecture, namely that the bank is not willing to exert monitoring effort when the loan is in a hopeless condition.

[Insert Table 1.6 here]

We have already pointed out that we focus on the effect of bank monitoring on loan repayment two quarters ahead, as this is the horizon in which we identify the strongest effect. Table 1.7 reports the estimates of various specifications of our 2SLS model for different horizons. The results show that the positive effect of bank monitoring on loan repayment is statistically significant only two quarters and three quarters ahead. Differently, if we rely on the Anderson-Rubin test to draw weak-identification-robust inference, we find that the effect of bank monitoring on loan repayment is significant for any horizon except for eight quarters ahead. Ignoring for a second the statistical significance and focusing exclusively on the sign and magnitude of the effect, we observe that the effect is already noticeable over the horizon of one quarter, it increases reaching is maximum two quarters ahead. Afterward, it slowly decreases until it disappears completely eight quarters ahead. This finding is consistent with the idea that a request for information may lead the bank to take specific actions to improve the likelihood that the firm repays its loan. For these actions to be effective it takes time and it seems that they deploy their effects mainly six months after the request takes place.

[Insert Table 1.7 here]

IRAP tax rate as a proxy for total bank monitoring

So far we have developed the empirical analysis using our measure for bank monitoring based on the requests for information made by banks on their existing borrowers. We have extensively discussed the limits of this measure, which is likely to grasp only a limited fraction of the actual intensity of bank monitoring. Relying on our theoretical model exposed in Section 1.2.3, we conjecture that variation in the IRAP tax rate affects any kind of bank monitoring activity. Thus, in order to capture the whole effect of bank monitoring on loan repayment, we directly mimic a change in total bank monitoring using the IRAP tax rate. Specifically, we run the following reduced form regression:

$$\text{NPL dummy}_{i,b,r,t} = \alpha + \beta \text{IRAP}_{r,t-6} + \boldsymbol{\gamma}'\boldsymbol{X}_{i,b,r,t-n} + \mu_{i,t} + \mu_b + \mu_r + \varepsilon_{i,b,r,t} \tag{1.6}$$

where $\boldsymbol{X}$ denotes the same vectors of loan, bank, and macro regional variables of the 2SLS model. We use the IRAP tax rate lagged of six quarters, as this tax is likely to exert its effect on bank monitoring only in the year in which the corresponding IRAP revenue is collected, as highlighted by Gambacorta et al. (2017). Moreover, to make sure that we are capturing only the effect of bank monitoring on loan repayment, as driven by the IRAP tax rate, we drop all observations in which we detect an increase in lending in the current quarter or in any of the previous quarters belonging to the year that follows that of the tax rate.[33]

Table 1.8 displays the results of this exercise. Looking at the first specification, we find that a decrease of 1 percentage point in the IRAP tax rate implies an increase of 4.9

[33]In fact, our model in Section 1.2.3, suggests that an decrease in the corporate tax rate implies a decrease in the lending rate. This, in turn, can cause an increase in credit demand from existing borrowers.

percentage points in the likelihood of loan distress. The magnitude of this effect is close but somewhat higher than that of the 2SLS model (3.9 percentage points). It is worth mentioning that the two models are run on different samples. However, the similarity in the magnitude of the effect provides us with an important insight on the relevance of our proxy for bank monitoring. If the requests for information were only partially correlated with other monitoring activities, the coefficient of the IRAP tax rate in this regression should have been sensibly higher than what observed in Table 1.8. The evidence that the magnitude of the effect of bank monitoring on loan repayment is similar and slightly stronger to what detected in the 2SLS model suggests that the requests for information are actually highly correlated with other forms of bank monitoring. In other words, the data is telling us that the bank's choice to monitor a borrower is strongly dichotomous. Either the bank does not monitor at all, or it carries out different kinds of monitoring activities when it decides to monitor. In fact, it is reasonable to think that, if a bank is concerned about the ability of a firm to meet its repayment schedule, it will control the condition of the other outstanding loans of the firm, check the financial reports, make visits on site and provide advisory services, all at the same time. As a consequence, despite our variable of bank monitoring captures only a fraction of the whole intensity of bank monitoring, it is able to grasp the effect of total bank monitoring on loan repayment.

When looking at the subsequent specifications, we observe that the pattern of the coefficients of *IRAP* resembles that of *Monitor* observed in Table 1.6. Also in this case, the positive effect of bank monitoring on loan repayment is mainly driven by the ability of the bank to prevent a given exposure from becoming unlikely-to-pay.

[Insert Table 1.8 here]

1.3.3 Robustness tests

It is reasonable to think that the performance condition of a loan is to some extent persistent over time. For example, if a loan is in distress it is likely that we it will still be so in the next quarter. Thus, as a first exercise, we check if our findings are robust to the inclusion of the dependent variable (a dummy equal to one if the loan is nonperforming and zero otherwise) lagged.[34] Table 1.9 reports the results of this robustness test. The first

[34] We include this variable lagged of three quarters so that it is predetermined with respect to the dependent variable in both stages of the 2SLS model.

stage of the 2SLS model is virtually unchanged. Interestingly, there is no evidence that banks monitor more intensively exposures that are already in a condition distress. This is hardly surprising though, as we highlighted that only a negligible fraction of requests for information (7.1%) are associated with nonperforming loans. The estimates of the second stage of the 2SLS model confirm that bank monitoring has a positive and significant effect on loan repayment, but this is somewhat weaker than what detected in Table 1.6 (albeit still economically strong). Specifically, an increase in the number of requests for information, that corresponds to a 1 percentage point decrease in the IRAP tax rate, implies a decrease by 2.8 percentage points in the likelihood that the loan ends up in distress two quarters ahead. Similarly, the coefficient of $IRAP$ in the reduced for model is still positive and statistically significant, but the magnitude is 1 percentage point lower than what observed in Table 1.8.

[Insert Table 1.9 here]

Despite only 7.1% of requests for information concern a credit exposure that is already in distress, we may wonder if our result is driven to some extent by loan restructuring rather than bank monitoring. In fact, if a loan gets in arrears the bank can loosen the borrower's financial constraints by postponing due payments or even providing additional credit (Brunner and Krahnen, 2008). Our data cleansing process implies that we already account for the former, as we dropped all relevant observations in which we detect an increase in lending to an existing borrower. Nevertheless, the bank may still respond to an overdue by modifying loan covenants. Thus, as a robustness test we further drop all observations pertaining to loans that are nonperforming at the time in which we observe bank monitoring, i.e. $t-2$. Table 1.10 reports the results of this exercise. We find that the negative effect of bank monitoring on loan repayment is still there, both in the 2SLS model and in the reduced form model, but the magnitude is somewhat smaller.

[Insert Table 1.10 here]

A third concern comes from the lags that we use for our independent variables in our main specifications. As we have already mentioned earlier, the lags of our regressors are defined in such way that each variable is predetermined with respect to the depend variable in each stage of our 2SLS model. This ensures that our 2SLS regression is not affected by issues of reverse causality. However, it exposes to the risk that the set of controls is not enough effective. In Table 1.11 we display the results of robustness checks in which we estimate our main regressions reducing the lag of the independent variables as much as possible. The results are virtually the same if compared to those of Table 1.6 and Table 1.8.

[Insert Table 1.11 here]

1.4 Conclusions

This paper investigates the effect of bank monitoring on loan repayment. Using granular loan-level information on business loans extended in Italy, we derive a novel proxy of bank monitoring. This consists in the number of requests for information made by banks on their existing borrowers to the Italian Credit Register.

To derive causal inference, we exploit taxation as a source of exogenous variation in bank monitoring. Our empirical strategy builds on a theoretical model that we develop to describe the effects of a corporate tax on bank monitoring incentives. The model predicts that a decrease in the corporate tax is associated with an increase in bank monitoring. This stems from two main channels. An increase in the corporate tax reduces bank profits and leads to a higher leverage ratio. These effects are only partially compensated by the pass-through of the corporate tax into the lending rate. As a result, an increase in the corporate tax implies a decrease in bank expected profits. This, in turn, weakens bank incentives to monitor borrowers. We use this theoretical prediction as the basis for our identification strategy.

Keeping in mind this theoretical prediction, we define our identification strategy using a 2SLS model in which bank monitoring is instrumented with a local tax rate (the Italy Regional Production Tax, IRAP, rate). To ensure that we control for potential confounding factors, we rely on firm-time fixed effects. These are aimed at capturing any time varying and time invariant, observable and unobservable condition of the firm that affects loan repayment. We find that an increase in the number of requests for information, as driven by a 1 percentage point decrease in the IRAP tax rate, reduces the probability of loan distress by 3.9 percentage points two quarters ahead.

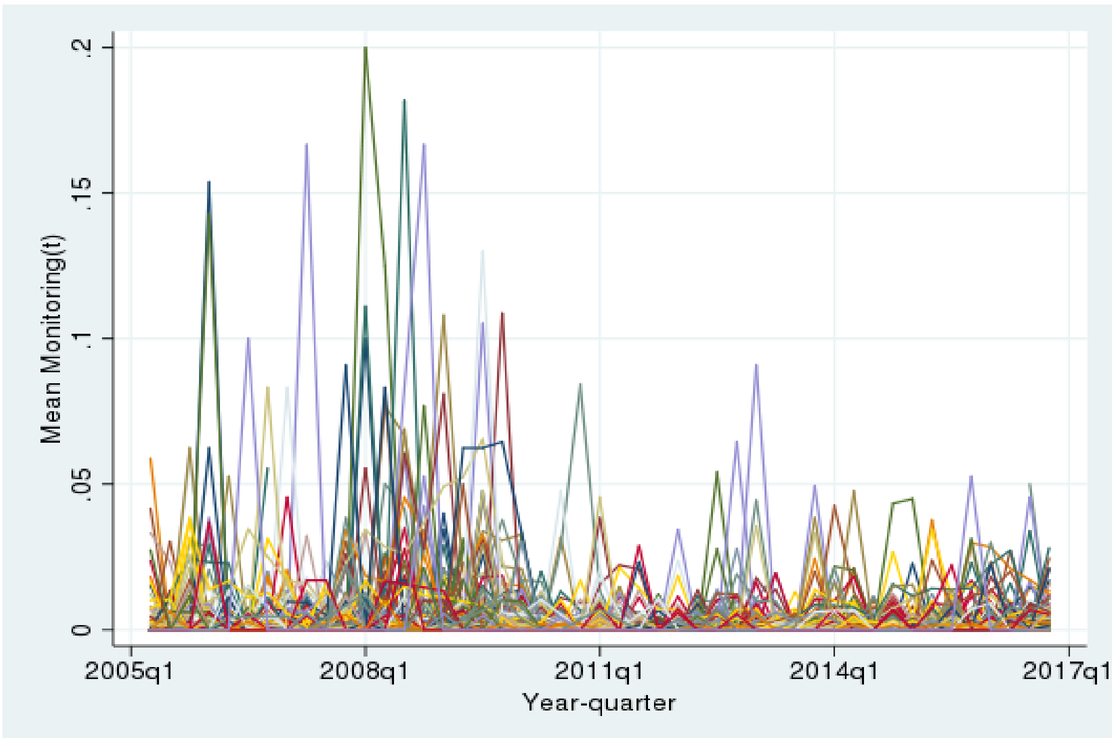
We acknowledge that our proxy of bank monitoring grasps only a fraction of the overall monitoring activity conducted by the bank. Thus, since the corporate tax rate is likely to influence bank incentives with respect to any form of monitoring, we extend our analysis estimating the effect of total bank monitoring, as driven by the IRAP tax rate, on the repayment performance of the firm. We find that this effect has a similar and somewhat higher magnitude to that exerted by the requests for information alone. We conclude that our proxy for bank monitoring is able to capture to a large extent the overall effect of bank monitoring on loan repayment.

Our findings have two key economic implications. First, the real effects of bank monitoring are substantial. Monitoring is valuable for individual banks, as it reduces default rates, as well as for the banking system as a whole. Second, taxation affects bank incentives to monitor borrowers in a significant way.

1.5 Figures

Figure 1.1: Bank requests for information over time

a. Average number of requests for information per borrower made by each bank



b. Average number of requests for information per borrower made by all banks

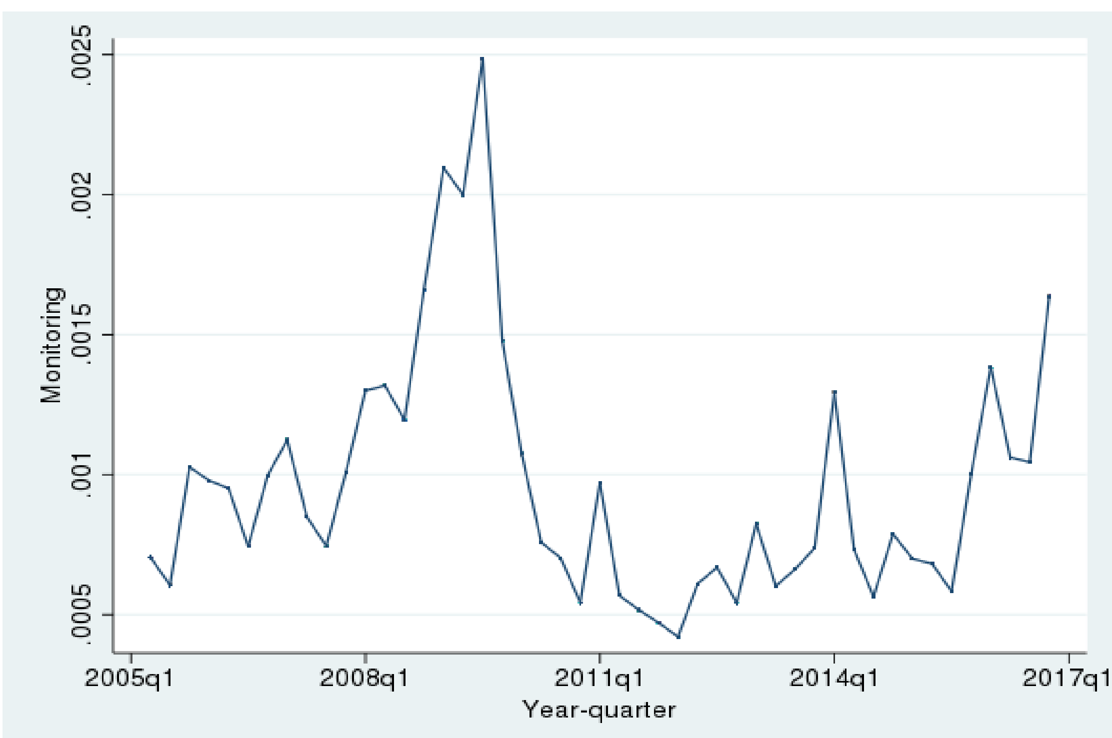


Figure 1.2: Bank requests for information associated with nonperforming loans and exposures close to the CR threshold

The figure shows the percentage of bank requests for information associated with nonperforming loans and each subcategory (blue bars), and the percentage of bank requests for information associated with credit exposures that are close to the thresholds to be included in the CR (green bar).

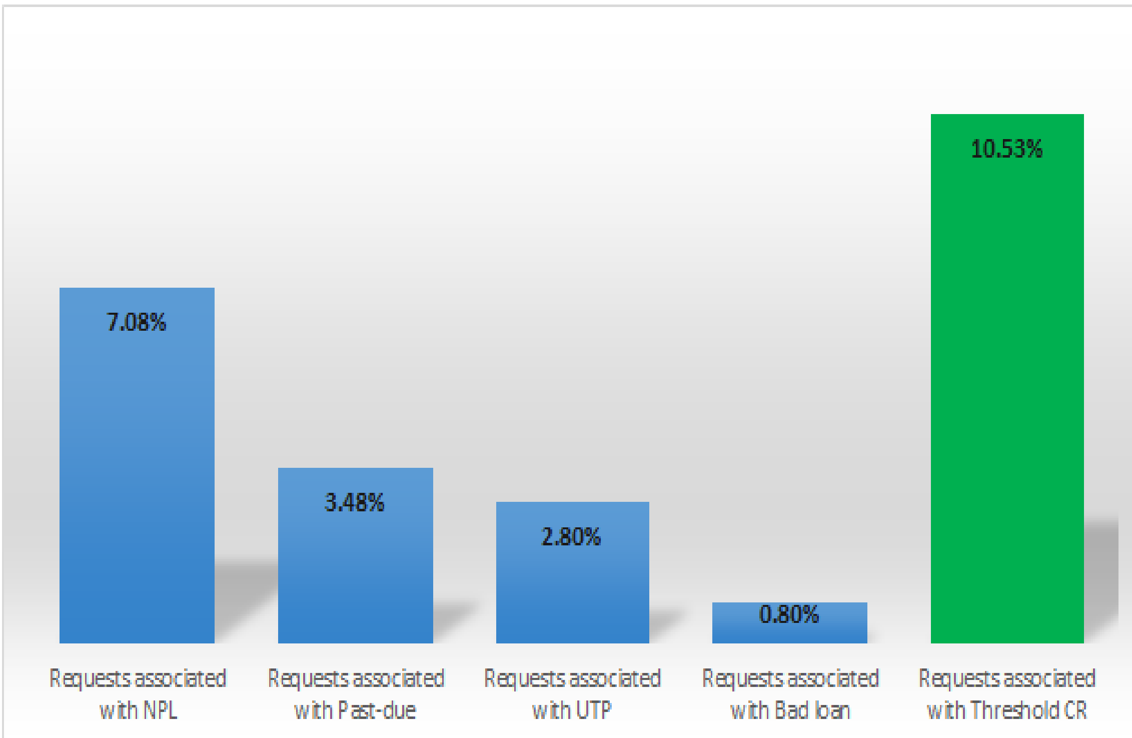
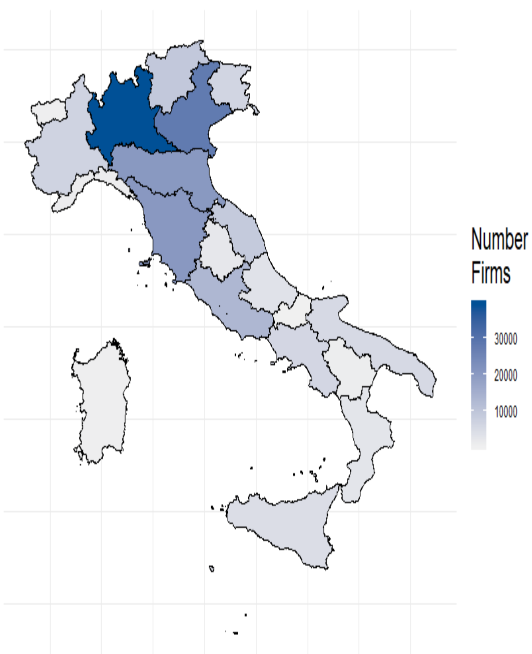


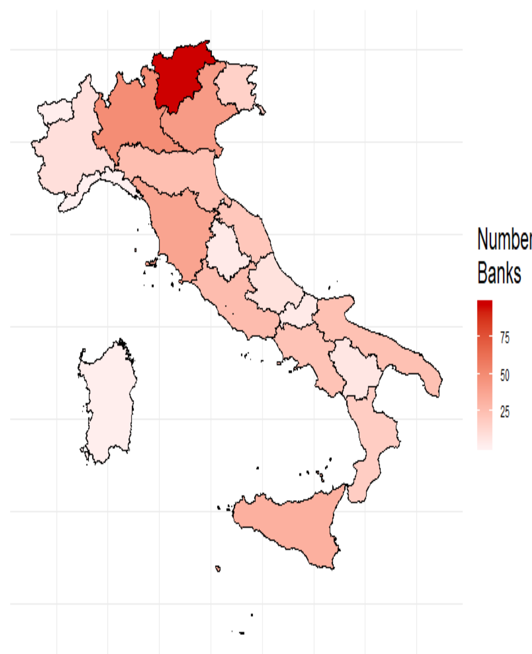
Figure 1.3: Distribution of firms and banks across Italian regions

The figure shows: the distribution of firms (a) and the distribution of banks (b) across Italian regions in the full sample; the distribution of firms (c) and the distribution of banks (d) across Italian regions in the reduced sample including only firms that have multiple credit relationships; the distribution of firms which borrow, at least in one quarter, from a bank located in a different region (e) and the distribution of banks which lend, at least in one quarter, to a firm located in a different region (f) across Italian regions in the reduced sample including only firms that have multiple credit relationships. The full sample includes 2,445,744 observations pertaining to 217,199 credit relationships having a duration greater than three quarters, and involving 176,781 firms and 446 banks over the period 2007-2016. The reduced sample has 556,717 observations and includes 53,776 credit relationships having a duration greater than three quarters, and involving 23,393 firms and 440 banks over the period 2007-2016.

a. Firms in the full sample



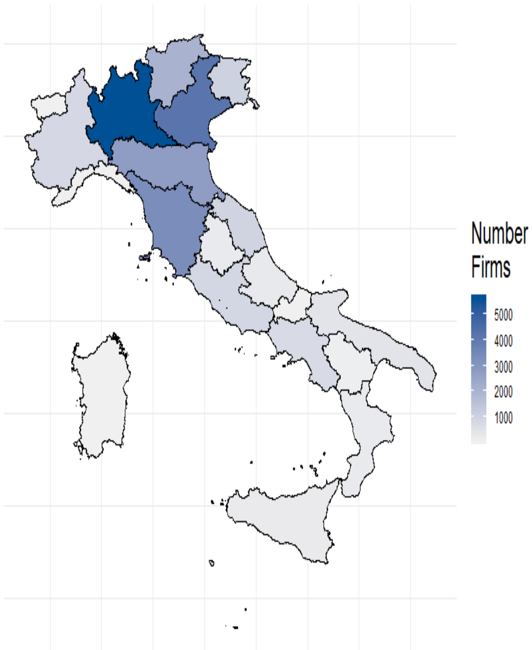
b. Banks in the full sample



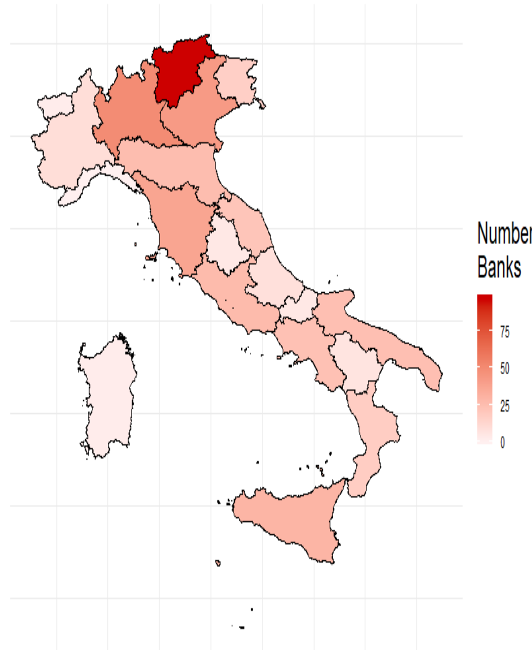
Continued on next page

Figure 1.3 – Continued from previous page

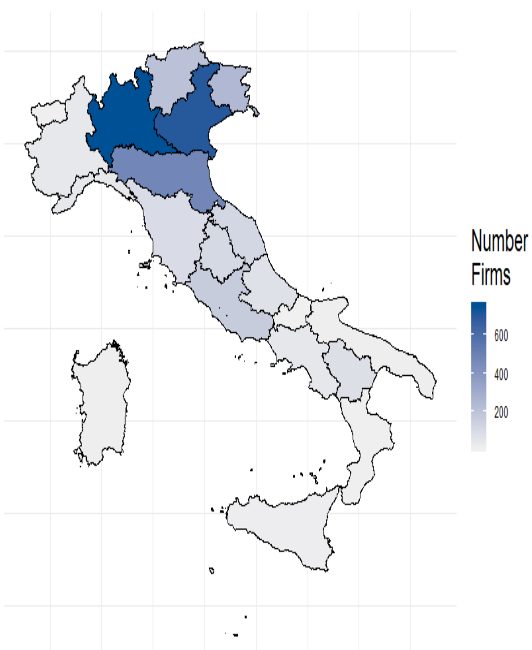
c. Firms in the reduced sample



d. Banks in the reduced sample



e. Firms which borrow from a bank located in a different region



f. Banks which lend to a firm located in a different region

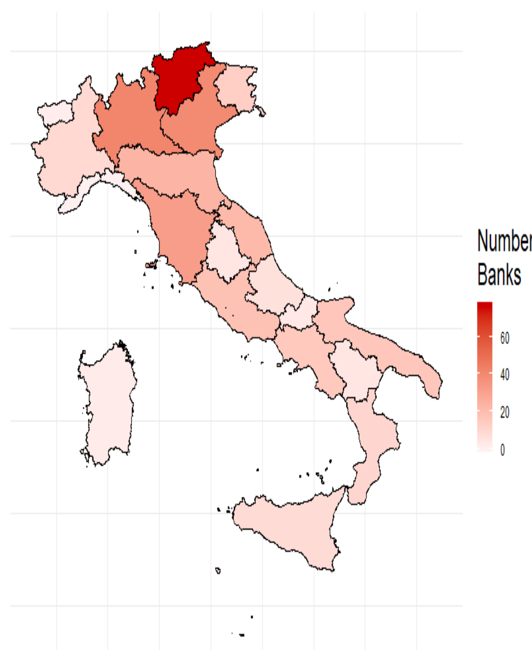
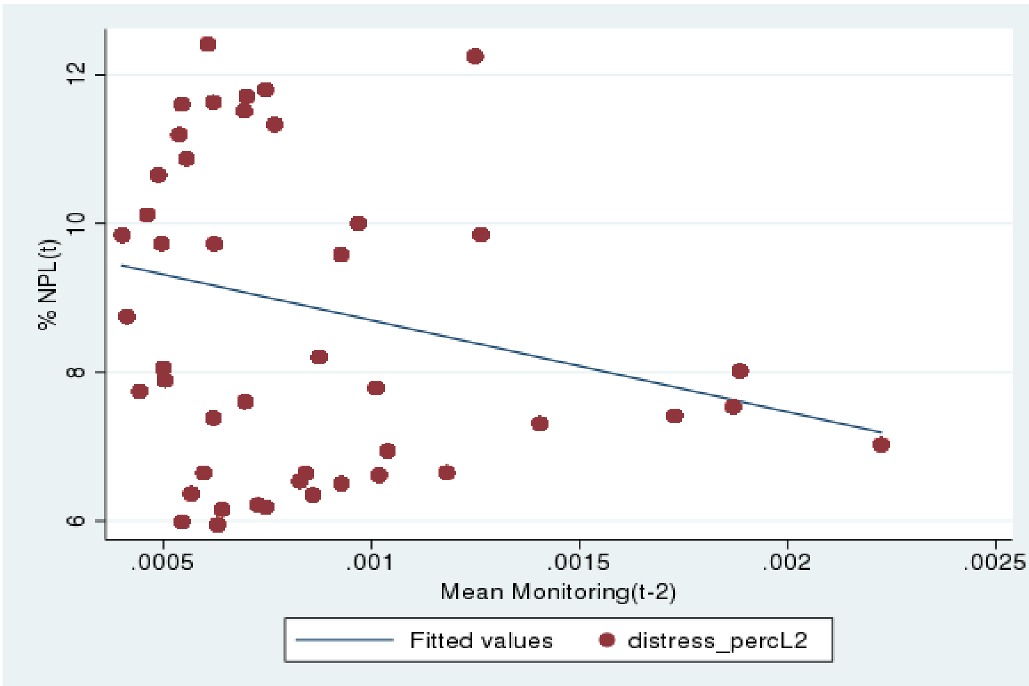


Figure 1.4: Nonperforming loans versus bank requests for information

The figure shows the percentage of nonperforming loans in each quarter (y-axis) against the average number of requests for information per loan submitted by banks two quarters before (x-axis). The plot refers to our baseline sample of 4,554,412 observations, covering the time period 2005-2016, from which we have dropped all observations in which we detect an increase in credit in the current quarter and/or in the previous two.



1.6 Tables

Table 1.1: Monitored firms and monitoring banks

The table reports panel regression estimates of different linear models investigating the relation between bank monitoring and a set of firm and bank variables. The dependent variable is displayed at the bottom of each column. The independent variables are described in Table A1.1 in Appendix B. For each independent variable the first row reports the coefficient, the second row reports in parenthesis the robust standard error that is corrected for multiclustering at the year-quarter and bank level. Fixed effects are either included, "Yes", not included, "No", spanned by another set of effects, "-", or not applicable, " ". ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. • denotes rescaled coefficients that have been multiplied by 10^4.

	Monitored firms			Monitoring banks
	(1)	(2)	(3)	(4)
	Monitor_t	Monitor_t	Monitor_t	Monitor_t
Share guarantee$_{t-1}$	-0.002**•	-0.002***•	-0.002***•	0.000
	(0.00)	(0.00)	(0.00)	(0.00)
Share exposure$_{t-1}$	-2.044***•	-2.212**•	-2.288**•	-4.497**•
	(0.00)	(0.00)	(0.00)	(0.00)
Length relation$_{t-1}$	-1.041***•	-1.293***•	-1.796***•	-0.991***•
	(0.00)	(0.00)	(0.00)	(0.00)
Size firm$_{t-4}$	0.566*•	-0.001***	-0.001***	
	(0.00)	(0.00)	(0.00)	
Size firm$_{t-4}$ $*$ Length relation$_{t-1}$			0.064***•	
			(0.00)	
ROA firm$_{t-4}$	0.000	0.001***	0.001***	
	(0.00)	(0.00)	(0.00)	
Credit score firm$_{t-4}$	0.418***•	-0.329*•	-0.329*•	
	(0.00)	(0.00)	(0.00)	
Capital ratio firm$_{t-4}$	0.000	0.000	0.000	
	(0.00)	(0.00)	(0.00)	
Capital ratio bank$_{t-4}$				-0.001
				(0.01)
ROA bank$_{t-4}$				0.020
				(0.02)
NPL ratio bank$_{t-4}$				-0.011
				(0.01)
Size bank$_{t-4}$				-0.002**
				(0.00)
Liquidity ratio bank$_{t-4}$				-0.016
				(0.02)
Nonretail deposit ratio bank$_{t-4}$				-0.004*
				(0.00)
GDP growth region firm$_{t-4}$	0.006	0.004	0.004	
	(0.00)	(0.00)	(0.00)	
Employment region firm$_{t-4}$	-0.000	0.000	0.000	
	(0.00)	(0.00)	(0.00)	
Inflation region firm$_{t-4}$	-0.000	0.000	0.000	
	(0.00)	(0.00)	(0.00)	
GDP growth region bank$_{t-4}$				-0.030

Continued on next page

				(0.02)
Employment region bank$_{t-4}$				0.000
				(0.00)
Inflation region bank$_{t-4}$				-0.000
				(0.00)
Year FE	-	-	-	-
Year-quarter FE	-	-	-	-
Region FE	Yes	Yes	Yes	-
Region bank FE	-	-	-	Yes
Bank FE	-	-	-	Yes
Firm FE	No	Yes	Yes	-
Industry firm FE	Yes	Yes	Yes	-
Firm-quarter FE	No	No	No	Yes
Bank-quarter FE	Yes	Yes	Yes	No
N	4201909	4195100	4195100	737713
R^2	0.012	0.113	0.113	0.469
Adjusted R^2	0.008	0.063	0.063	0.008
F-test statistic	9.852***	14.083***	13.066***	7.803***
degrees of freedom	(10, 46)	(10, 46)	(11, 46)	(12, 39)

Table 1.2: IRAP tax rates

The table reports the basic national IRAP tax rate and the regional IRAP tax rates applied to banks during 2006-2016.

	IRAP tax rate (%)										
	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Basic National IRAP	4.25	4.25	3.90	3.90	3.90	4.65	4.65	4.65	4.65	4.65	4.65
Region											
Abruzzo	5.25	5.25	4.82	4.82	4.82	5.57	5.57	5.57	5.57	5.57	5.57
Basilicata	4.25	4.25	3.90	3.90	3.90	4.65	4.65	4.65	4.65	4.65	4.65
Calabria	4.25	4.25	3.90	4.82	4.97	5.72	5.72	5.72	5.57	5.57	5.57
Campania	5.25	5.25	4.82	4.82	4.97	5.72	5.72	5.72	5.72	5.72	5.72
Emilia-Romagna	4.25	5.25	4.82	4.82	4.82	5.57	5.57	5.57	5.57	5.57	5.57
Friuli-Venezia Giulia	4.25	4.25	3.90	3.90	3.90	4.65	4.65	4.65	4.65	4.65	4.65
Lazio	5.25	5.25	4.82	4.82	4.97	5.57	5.57	5.57	5.57	5.57	5.57
Liguria	5.25	5.25	4.82	4.82	4.82	5.57	5.57	5.57	5.57	5.57	5.57
Lombardia	5.25	5.25	4.82	4.82	4.82	5.57	5.57	5.57	5.57	5.57	5.57
Marche	5.15	5.15	4.73	4.73	4.73	5.48	5.48	5.48	5.48	5.48	5.48
Molise	5.25	5.25	4.82	4.82	4.97	5.72	5.72	5.72	5.72	5.72	5.57
Piemonte	4.25	5.25	4.82	4.82	4.82	5.57	5.57	5.57	5.57	5.57	5.57
Puglia	4.25	4.25	4.82	4.82	4.82	5.57	5.57	5.57	5.57	5.57	5.57
Sardegna	4.25	4.25	3.90	3.90	3.90	4.65	4.65	1.4	1.4	5.57	5.57
Sicilia	5.25	5.25	4.82	4.82	4.82	5.57	5.57	5.57	5.57	5.57	5.57
Toscana	4.40	5.25	4.82	4.82	4.82	5.57	5.57	5.57	5.57	5.57	5.57
Trentino-Alto Adige	4.25	4.25	3.44	3.40	3.19	4.65	4.45	4.45	4.65	4.65	4.65
Umbria	4.25	4.25	4.82	4.82	4.82	5.57	5.57	5.57	5.57	5.57	5.57
Val D'Aosta	4.25	4.25	3.90	3.90	3.90	4.65	4.65	4.65	4.65	4.65	4.65
Veneto	5.25	5.25	4.82	4.82	4.82	5.57	5.57	5.57	5.57	5.57	5.57
Minimum	4.25	4.25	3.44	3.40	3.19	4.65	4.45	1.40	1.40	4.65	4.65
Mean	4.70	4.85	4.52	4.56	4.58	5.36	5.35	5.19	5.19	5.40	5.39
Maximum	5.25	5.25	4.82	4.82	4.97	5.72	5.72	5.72	5.72	5.72	5.72

Table 1.3: Descriptive statistics

The table reports summary statistics for the variables used in our regression analysis. In panel A, statistics refer to our full sample of 2,445,744 observations. This sample includes 217,199 credit relationships having a duration greater than three quarters, and involving 176,781 firms and 446 banks over the period 2007-2016. In panel B, statistics refer to our reduced sample including only firms that have multiple credit relationships. This panel has 556,717 observations and includes 53,776 credit relationships having a duration greater than three quarters, and involving 23,393 firms and 440 banks over the period 2007-2016.

Panel A: Full sample						
Variable Name	Obs.	Mean	St. Dev.	Minimum	Median	Maximum
Loan-level Variables						
NPL dummy$_t$	2,445,744	0.082	0.274	0.000	0.000	1.000
Past-due dummy$_t$	2,445,744	0.024	0.153	0.000	0.000	1.000
UTP dummy$_t$	2,445,744	0.037	0.189	0.000	0.000	1.000
Bad loan dummy$_t$	2,445,744	0.021	0.143	0.000	0.000	1.000
Monitor$_{t-2}$	2,445,744	0.001	0.030	0.000	0.000	4.000
Length relation$_{t-2}$	2,445,744	17.873	10.971	2.000	16.000	46.000
Share guarantee$_{t-2}$	2,445,744	0.048	0.542	0.000	0.000	17.374
Share exposure$_{t-2}$	2,445,744	0.414	0.390	0.000	0.265	1.000
Firm-level Variables						
Size firm$_{t-5}$	2,445,744	7.359	1.420	0.000	7.265	18.024
ROA firm$_{t-5}$	2,445,744	-0.004	0.104	-3.167	0.003	0.411
Credit score firm$_{t-5}$	2,445,744	5.182	1.887	1.000	5.000	9.000
Capital ratio firm$_{t-5}$	2,445,744	0.173	3.745	-5804.691	0.137	1.000
Bank-level Variables						
Size bank$_{t-5}$	2,445,744	6.723	0.975	1.557	6.706	9.351
Capital ratio bank$_{t-5}$	2,445,744	0.087	0.025	0.024	0.084	0.682
Liquidity ratio bank$_{t-5}$	2,445,744	0.005	0.003	0.001	0.004	0.044
ROA bank$_{t-5}$	2,445,744	0.003	0.006	-0.056	0.003	0.043
Nonretail deposit ratio bank$_{t-5}$	2,445,744	0.170	0.070	0.005	0.164	0.715
NPL ratio bank$_{t-5}$	2,445,744	0.115	0.063	0.021	0.102	0.415
Regional Variables						
IRAP$_{t-5}$	2,445,744	0.051	0.005	0.014	0.053	0.057
GDP growth region firm$_{t-5}$	2,445,744	-0.004	0.027	-0.088	0.004	0.085
Employment region firm$_{t-5}$	2,445,744	48.286	5.251	30.350	49.900	55.200
Inflation region firm$_{t-5}$	2,445,744	1.621	1.115	-0.200	1.500	4.400
GDP growth region bank$_{t-5}$	2,445,744	-0.004	0.026	-0.088	0.004	0.085
Employment region bank$_{t-5}$	2,445,744	48.371	5.254	30.350	49.900	55.200
Inflation region bank$_{t-5}$	2,445,744	1.622	1.114	-0.200	1.500	4.400
Panel B: Multiple bank relationships						
Loan-level Variables						
NPL dummy$_t$	556,717	0.114	0.317	0.000	0.000	1.000
Past-due dummy$_t$	556,717	0.023	0.151	0.000	0.000	1.000
UTP dummy$_t$	556,717	0.050	0.219	0.000	0.000	1.000
Bad loan dummy$_t$	556,717	0.040	0.196	0.000	0.000	1.000

Continued on next page

Table 1.3 – *Continued from previous page*

Monitor$_{t-2}$	556,717	0.001	0.029	0.000	0.000	2.000
Length relation$_{t-3}$	556,717	17.863	11.039	1.000	16.000	45.000
Share guarantee$_{t-3}$	556,717	6.868	1027.814	0.000	0.000	537526.100
Share exposure$_{t-3}$	556,717	0.215	0.238	0.000	0.126	1.000
Firm-level Variables						
Size firm$_{t-6}$	296,828	8.151	1.403	0.000	8.050	15.665
ROA firm$_{t-6}$	295,690	-0.009	0.105	-1.787	0.002	0.407
Credit score firm$_{t-6}$	293,719	5.444	1.757	1.000	6.000	9.000
Capital ratio firm$_{t-6}$	296,828	0.104	1.529	-200.500	0.123	1.000
Bank-level Variables						
Size bank$_{t-6}$	556,717	6.672	0.881	2.420	6.683	9.351
Capital ratio bank$_{t-6}$	556,717	0.089	0.025	0.024	0.085	0.405
Liquidity ratio bank$_{t-6}$	556,717	0.005	0.003	0.001	0.004	0.044
ROA bank$_{t-6}$	556,717	0.003	0.005	-0.051	0.003	0.024
Nonretail deposit ratio bank$_{t-6}$	556,717	0.168	0.066	0.005	0.161	0.715
NPL ratio bank$_{t-6}$	556,717	0.099	0.055	0.021	0.083	0.367
Regional Variables						
IRAP$_{t-6}$	556,717	0.050	0.006	0.014	0.053	0.057
GDP growth region firm$_{t-6}$	551,028	-0.004	0.027	-0.088	0.005	0.085
Employment region firm$_{t-6}$	551,028	49.470	4.491	30.350	50.200	55.200
Inflation region firm$_{t-6}$	551,026	1.723	1.088	-0.200	1.600	4.400
GDP growth region bank$_{t-6}$	556,717	-0.004	0.027	-0.088	0.005	0.085
Employment region bank$_{t-6}$	556,717	49.577	4.515	30.350	50.200	55.200
Inflation region bank$_{t-6}$	556,717	1.730	1.086	-0.200	1.600	4.400

Table 1.4: Preliminary analysis on bank monitoring and loan repayment

The table reports panel regression estimates of a linear probability model analyzing the relation between bank monitoring and (i) loan repayment (specifications (1)-(4) and (7)), and (ii) firm's conditions (specifications (5)-(6)). The dependent variable is displayed at the bottom of each column. The independent variables are described in Table A1.1 in Appendix B. For each independent variable the first row reports the coefficient, the second row reports in parenthesis the robust standard error that is corrected for multiclustering at the year-quarter and bank level. Fixed effects are either included, "Yes", not included, "No", or spanned by another set of effects, "-". ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. • denotes rescaled coefficients that have been multiplied by 10^2.

	(1) NPL dummy$_t$	(2) Past-due dummy$_t$	(3) UTP dummy$_t$	(4) Bad loan dummy$_t$	(5) Capital ratio firm$_t$	(6) ROA firm$_t$	(7) Credit score$_t$	(8) NPL dummy$_t$
Monitor$_{t-2}$	-0.008***	0.004	-0.002	-0.010***	0.013**	0.002	-0.037	0.027*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.01)
Share guarantee$_{t-2}$	-0.000	-0.000	0.001	-0.001*	0.002**	-0.000	-0.006**	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Share exposure$_{t-2}$	0.011***	0.004***	0.008***	-0.001	-0.010***	-0.000	0.122***	0.027***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Length relation$_{t-2}$	0.001***	0.009**•	0.001***	-0.000	0.026***•	-0.000	-0.002***	0.001***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Size firm$_{t-5}$	-0.049***	0.002***	-0.008***	-0.043***	0.026***	0.003***	-0.013	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	
ROA firm$_{t-5}$	-0.267***	-0.019***	-0.101***	-0.147***	0.360***		-2.174***	
	(0.01)	(0.00)	(0.01)	(0.02)	(0.01)		(0.15)	
Credit score firm$_{t-5}$	0.015***	0.001***	0.006***	0.007***	-0.032***	-0.004***		
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
Capital ratio firm$_{t-5}$	-0.000	0.000	0.000	-0.000		-0.000	-0.000	
	(0.00)	(0.00)	(0.00)	(0.00)		(0.00)	(0.00)	
Capital ratio bank$_{t-5}$	-0.220**	0.124***	-0.155**	-0.189***	0.169***	0.021	-0.263	0.113
	(0.10)	(0.04)	(0.08)	(0.05)	(0.05)	(0.02)	(0.38)	(0.08)
ROA bank$_{t-5}$	-0.823***	0.143*	-0.574***	-0.391***	0.224***	0.069**	-0.338	-0.451***
	(0.16)	(0.07)	(0.12)	(0.09)	(0.06)	(0.03)	(0.45)	(0.17)
NPL ratio bank$_{t-5}$	0.042	-0.010	-0.007	0.060*	-0.126***	-0.007	-0.128	-0.056
	(0.06)	(0.02)	(0.04)	(0.03)	(0.04)	(0.01)	(0.27)	(0.09)
Size bank$_{t-5}$	-0.003	0.008**	-0.002	-0.009**	0.000	-0.003**	0.027	0.012*
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.03)	(0.01)
Liquidity ratio bank$_{t-5}$	-0.050	-0.052	-0.102	0.102	0.050	0.096**	-2.440	0.399
	(0.17)	(0.13)	(0.11)	(0.20)	(0.09)	(0.04)	(2.11)	(0.24)
Nonretail deposit ratio$_{t-5}$	-0.022	0.015	-0.020	-0.017*	-0.003	0.010*	-0.232**	0.042*
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.10)	(0.02)
GDP growth region firm$_{t-5}$	0.005	-0.079*	0.047	0.035	0.026	0.042*	-0.439	
	(0.07)	(0.04)	(0.05)	(0.04)	(0.03)	(0.02)	(0.32)	
Employment region firm$_{t-5}$	-0.008***	0.001	-0.006***	-0.004**	0.001	-0.000	0.009	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	
Inflation region firm$_{t-5}$	-0.005	-0.002	-0.002	-0.002	0.001	0.001	-0.008	
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	
GDP growth region bank$_{t-5}$	0.000	0.073	-0.054	-0.017	-0.007	-0.048*	0.211	-0.175
	(0.07)	(0.05)	(0.06)	(0.04)	(0.04)	(0.02)	(0.34)	(0.14)
Employment region bank$_{t-5}$	0.007**	-0.001	0.006***	0.003	-0.000	-0.000	-0.004	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Inflation region bank$_{t-5}$	0.006	0.001	0.002	0.003	-0.003	-0.000	-0.005	0.011
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.01)
Year-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-
Region bank FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-
Industry firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-
Firm-quarter FE	No	No	No	No	No	No	No	Yes
N	2445744	2445744	2445744	2445744	2426489	2404771	2431398	564109
R^2	0.566	0.254	0.442	0.572	0.818	0.519	0.794	0.836
Adjusted R^2	0.533	0.196	0.398	0.538	0.804	0.481	0.777	0.693
F-test statistic	85.167***	17.219***	49.160***	30.669***	161.154***	5.684***	16.070***	16.372***
degrees of freedom	(20, 38)	(20, 38)	(20, 38)	(20, 38)	(19, 38)	(19, 38)	(19, 38)	(13, 38)

Table 1.5: Exogeneity analysis of the IRAP tax rate

The table reports panel regression estimates of a linear model analyzing the effect of different variables on the IRAP tax rate. *Basic IRAP* is the basic IRAP tax rate for banks defined at the national level. Δ*IRAP health* is a dummy equal to one if an increase in the IRAP tax rate occurs in response to a regional health deficit. *Capital ratio region bank* and *ROA region bank* are the aggregate capital ratio and ROA of the banking system at the regional level, respectively. *NPL ratio region bank* is the average ratio of nonperforming loans of banks operating in a specific region. The remaining variables are described in Table A1.1 in Appendix B. The sample of specifications (1)-(2) covers the time period 2005-2016. The sample of specifications (3)-(6) covers instead the time period 2007-2016, because of a lack in the availability of data on bank conditions before 2007. The dependent variable is displayed at the bottom of each column. For each independent variable the first row reports the coefficient, the second row reports in parenthesis the robust standard error that is corrected for multiclustering at the year and regional level. Fixed effects are included, "Yes". ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

	(1) $IRAP_t$	(2) $IRAP_t$	(3) $IRAP_t$	(4) $IRAP_t$	(5) $IRAP_t$	(6) $IRAP_t$
GDP growth region $bank_{t-1}$	0.101	0.150			0.535	0.535
	(1.41)	(1.41)			(0.82)	(0.81)
Employment region $bank_{t-1}$	-0.027	-0.026			-0.011	-0.010
	(0.02)	(0.02)			(0.01)	(0.01)
Inflation region $bank_{t-1}$	-0.030	-0.028			-0.026	-0.023
	(0.02)	(0.03)			(0.02)	(0.02)
Basic $IRAP_t$	0.976***	0.990***	1.049***	1.066***	1.023***	1.035***
	(0.11)	(0.11)	(0.04)	(0.03)	(0.05)	(0.04)
ΔIRAP $health_t$		0.131**		0.138**		0.104
		(0.05)		(0.05)		(0.06)
Capital ratio region $bank_{t-1}$			2.144	3.894	3.285	3.535
			(5.18)	(5.66)	(4.91)	(4.91)
ROA region $bank_{t-1}$			-2.891	-0.918	-2.080	-1.748
			(8.67)	(8.75)	(8.36)	(8.01)
NPL ratio region $bank_{t-1}$			0.005	0.006	-0.001	0.000
			(0.01)	(0.01)	(0.01)	(0.01)
Region bank FE	Yes	Yes	Yes	Yes	Yes	Yes
N	228	228	190	190	186	186
R^2	0.6656	0.6666	0.7301	0.7287	0.7223	0.7227
Adjusted R^2	0.6279	0.6272	0.6927	0.6911	0.6789	0.6774

Table 1.6: 2SLS model

The table reports panel regression estimates of the 2SLS model of equation 1.5 analyzing the effect of bank monitoring on loan repayment. In this model we have one endogenous variable, *Monitor*, which is instrumented with the IRAP tax rate, *IRAP*. The dependent variable is displayed at the bottom of each column. The independent variables are described in Table A1.1 in Appendix B. For each independent variable the first row reports the coefficient, the second row reports in parenthesis the robust standard error that is corrected for multiclustering at the year-quarter and bank level. Fixed effects are included, "Yes". ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. • denotes rescaled coefficients that have been multiplied by 10, whereas •• denotes rescaled coefficients that have been multiplied by 10^6. At the bottom of the table we report the results of standard underidentification and weak identification tests, as well as the Anderson-Rubin test.

	First stage	Second stage			
	(1)	(2)	(3)	(4)	(5)
	Monitor_{t-2}	NPL dummy_t	Past-due dummy_t	UTP dummy_t	Bad loan dummy_t
Monitor_{t-2}		-9.425*	-4.727	-5.761*	0.854
		(5.24)	(3.89)	(3.34)	(0.89)
IRAP_{t-6}	-0.413**				
	(0.19)				
Share guarantee_{t-3}	-0.005*••	0.000	0.000	-0.000	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Share exposure_{t-3}	-0.000	0.025***	0.011***	0.015***	-0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Length relation_{t-3}	-0.001***•	-0.000	-0.000	-0.000	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Capital ratio bank_{t-6}	0.019	0.344	0.282**	0.087	-0.024
	(0.02)	(0.21)	(0.13)	(0.13)	(0.06)
ROA bankt-6	-0.001	-0.360	0.131	-0.298	-0.192**
	(0.02)	(0.32)	(0.17)	(0.20)	(0.09)
NPL ratio bank_{t-6}	-0.005	-0.077	0.022	0.014	-0.112*
	(0.02)	(0.18)	(0.10)	(0.12)	(0.06)
Size bankt-6	-0.001	0.003	-0.004	0.009	-0.004
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)
Liquidity ratio bank_{t-6}	0.008	0.453	0.140	0.160	0.152
	(0.04)	(0.40)	(0.26)	(0.30)	(0.11)
Nonretail deposit ratio bank_{t-6}	-0.004	0.002	-0.010	0.010	0.002
	(0.00)	(0.03)	(0.02)	(0.02)	(0.01)
GDP growth region bank_{t-6}	-0.048	-0.500*	-0.341*	-0.215	0.049
	(0.03)	(0.29)	(0.19)	(0.20)	(0.06)
Employment region bank_{t-6}	0.001*	0.006	0.004	0.004	-0.002
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
Inflation region bank_{t-6}	-0.001	0.004	-0.009	0.019**	-0.006
	(0.00)	(0.02)	(0.01)	(0.01)	(0.00)
Firm-quarter FE	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes
Region bank FE	Yes	Yes	Yes	Yes	Yes

Continued on next page

N	711901	556717	556717	556717	556717
R^2	0.469				
Adjusted R^2	0.005				
F-test statistic	4.75***	8.980***	3.714***	6.989***	1.905**
degrees of freedom	(13, 37)	(13, 37)	(13, 37)	(13, 37)	(13, 37)
Underidentification test					
Kleibergen-Paap LM statistic	4.00				
Chi-sq P-val	0.046				
Weak identification test					
Kleibergen-Paap Wald F-stat	4.94				
Cragg-Donald Wald F-stat	18.11				
Stock-Yogo critical value 10% maximal IV size	16.38				
			Anderson-Rubin test		
Anderson-Rubin Wald statistic		6.53**	1.90	5.87**	1.08
Chi-sq P-val		0.011	0.168	0.015	0.300

Table 1.7: 2SLS model with different horizons

The table reports panel regression estimates of the 2SLS model analyzing the effect of bank monitoring on loan repayment for different horizons. In this model we have one endogenous variable, *Monitor*, which is instrumented with the IRAP tax rate, *IRAP*. The dependent variable is displayed at the bottom of each column. The independent variables are described in Table A1.1 in Appendix B. For each independent variable the first row reports the coefficient, the second row reports in parenthesis the robust standard error that is corrected for multiclustering at the year-quarter and bank level. Fixed effects are included, "Yes". ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. • denotes rescaled coefficients that have been multiplied by 10, whereas •• denotes rescaled coefficients that have been multiplied by 10^6. At the bottom of the table we report the results of standard underidentification and weak identification tests, as well as the Anderson-Rubin test.

	One quarter		Two quarters		Three quarters		Four quarters		Eight quarters	
	First stage	Second stage	First stage	Second stage	First stage	Second stage	First stage	Second stage	First stage	Second stage
	(1) Monitor_{t-1}	(2) NPL dummy_t	(3) Monitor_{t-2}	(4) NPL dummy_t	(5) Monitor_{t-3}	(6) NPL dummy_t	(7) Monitor_{t-4}	(8) NPL dummy_t	(9) Monitor_{t-8}	(10) NPL dummy_t
Monitor_{t-n}		-10.805		-9.425*		-6.979*		-6.263		-6.804
		(7.44)		(5.24)		(3.99)		(3.89)		(6.22)
IRAP_{t-n-4}	-0.319**		-0.413**		-0.448**		-0.475**		-0.003	
	(0.16)		(0.19)		(0.21)		(0.22)		(0.00)	
$\text{Share guarantee}_{t-n-1}$	0.000	-0.000	-0.005*••	0.000	-0.000	0.000	-0.000	0.000	0.000	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\text{Share exposure}_{t-n-1}$	-0.004*•	0.021***	-0.000	0.025***	-0.000	0.028***	-0.000	0.028***	-0.000	0.026***
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\text{Length relation}_{t-n-1}$	-0.001***•	-0.001	-0.001***•	-0.000	-0.001***•	-0.000	-0.001***•	0.000	-0.001***•	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\text{Capital ratio bank}_{t-n-6}$	0.012	0.224	0.019	0.344	0.014	0.326**	0.034*	0.460**	0.067	0.727
	(0.02)	(0.20)	(0.02)	(0.21)	(0.02)	(0.16)	(0.02)	(0.19)	(0.04)	(0.50)
ROA bank_{t-n-6}	0.002	-0.441	-0.001	-0.360	-0.007	-0.198	0.008	-0.161	0.008	-0.182
	(0.02)	(0.32)	(0.02)	(0.32)	(0.02)	(0.27)	(0.02)	(0.25)	(0.04)	(0.37)
$\text{NPL ratio bank}_{t-n-6}$	-0.013	-0.178	-0.005	-0.077	-0.017	-0.193	-0.007	-0.104	-0.047	-0.476
	(0.01)	(0.20)	(0.02)	(0.18)	(0.02)	(0.18)	(0.02)	(0.15)	(0.04)	(0.35)
Size bank_{t-n-6}	-0.001*	-0.004	-0.001	0.003	-0.001	0.006	-0.000	0.009	0.001	0.028*
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
$\text{Liquidity ratio bank}_{t-n-6}$	-0.022	0.043	0.008	0.453	0.039	0.427	0.048	0.423	0.081	0.913
	(0.03)	(0.40)	(0.04)	(0.40)	(0.04)	(0.43)	(0.05)	(0.47)	(0.09)	(0.77)
$\text{Nonretail deposit ratio bank}_{t-n-6}$	-0.005**	-0.017	-0.004	0.002	-0.003	0.016	-0.003	0.001	-0.000	0.031
	(0.00)	(0.05)	(0.00)	(0.03)	(0.00)	(0.03)	(0.00)	(0.03)	(0.00)	(0.04)
$\text{GDP growth region bank}_{t-n-6}$	-0.058**	-0.788**	-0.048	-0.500*	-0.051	-0.386	-0.057	-0.158	-0.082*	-0.211
	(0.03)	(0.38)	(0.03)	(0.29)	(0.04)	(0.29)	(0.04)	(0.34)	(0.04)	(0.58)
$\text{Employment region bank}_{t-n-6}$	0.001	0.004	0.001*	0.006	0.001**	0.010	0.001*	0.012*	0.001	0.006
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
$\text{Inflation region bank}_{t-n-6}$	-0.001	-0.001	-0.001	0.004	-0.002	-0.005	-0.002	-0.012	-0.001	-0.003
	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.02)
Firm-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region bank FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	676180	676180	556717	556717	461912	461912	389378	389378	205002	205002
R^2	0.470		0.469		0.473		0.472		0.470	
Adjusted R^2	0.009		0.005		0.010		0.006		-0.005	
F-test statistic	6.10***	7.413***	4.75***	8.980***	3.84***	12.729***	3.11***	12.747***	6.25***	6.085***
degrees of freedom	(13, 38)	(13, 38)	(13, 37)	(13, 37)	(13, 36)	(13, 36)	(13, 35)	(13, 35)	(13, 31)	(13, 31)
Change in the likelihood of loan distress (in p.p.) in response to an increase in the number of requests for information that correspond to a decrease of 1 p.p. in the tax rate		-3.45		-3.89*		-3.13*		-2.97		-0.02
Underidentification test										
Kleibergen-Paap LM statistic	3.56		4.00		3.72		3.86		2.23	
Chi-sq P-val	0.059		0.046		0.054		0.050		0.135	
Weak identification test										
Kleibergen-Paap Wald F statistic	4.13		4.94		4.59		4.65		2.33	
Cragg-Donald Wald F statistic	11.80		18.11		19.56		19.06		4.89	
Stock-Yogo critical value 10% maximal IV size	16.38		16.38		16.38		16.38		16.38	
Anderson-Rubin test										
Anderson-Rubin Wald statistic		5.43**		6.53**		4.76**		4.27**		2.11
Chi-sq P-val		0.020		0.011		0.029		0.039		0.146

Table 1.8: Reduced form model

The table reports panel regression estimates of the reduced form linear model of equation 1.6 analyzing the effect of bank monitoring, as driven by the IRAP tax rate, on loan repayment. The dependent variable is displayed at the bottom of each column. The independent variables are described in Table A1.1 in Appendix B. For each independent variable the first row reports the coefficient, the second row reports in parenthesis the robust standard error that is corrected for multiclustering at the year-quarter and bank level. Fixed effects are included, "Yes". ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. • denotes rescaled coefficients that have been multiplied by 10^3.

	(1) NPL dummy$_t$	(2) Past-due dummy$_t$	(3) UTP dummy$_t$	(4) Bad loan dummy$_t$
IRAP$_{t-6}$	4.873***	1.718	3.694***	-0.509
	(1.79)	(1.61)	(1.07)	(0.39)
Share guarantee$_{t-3}$	0.000	0.001*•	-0.000	0.000
	(0.00)	(0.00)	(0.00)	(0.00)
Share exposure$_{t-3}$	0.027***	0.012***	0.017***	-0.002
	(0.00)	(0.00)	(0.00)	(0.00)
Length relation$_{t-3}$	0.001***	-0.000	0.001***	0.000
	(0.00)	(0.00)	(0.00)	(0.00)
Capital ratio bank$_{t-6}$	0.210**	0.215***	-0.007	0.000
	(0.09)	(0.06)	(0.10)	(0.06)
ROA bank$_{t-6}$	-0.318	0.100	-0.207	-0.207*
	(0.20)	(0.11)	(0.17)	(0.11)
NPL ratio bank$_{t-6}$	-0.053	0.040	0.057	-0.146**
	(0.11)	(0.08)	(0.10)	(0.07)
Size bank$_{t-6}$	0.013*	0.002	0.016**	-0.005
	(0.01)	(0.00)	(0.01)	(0.00)
Liquidity ratio bank$_{t-6}$	0.531**	0.085	0.206	0.238*
	(0.25)	(0.22)	(0.22)	(0.13)
Nonretail deposit ratio bank$_{t-6}$	0.049**	0.021	0.032*	-0.004
	(0.02)	(0.01)	(0.02)	(0.01)
GDP growth region bank$_{t-6}$	-0.053	-0.139	0.082	0.017
	(0.17)	(0.12)	(0.15)	(0.06)
Employment region bank$_{t-6}$	-0.000	0.001	0.000	-0.002
	(0.01)	(0.00)	(0.00)	(0.00)
Inflation region bank$_{t-6}$	0.024**	-0.002	0.032***	-0.006
	(0.01)	(0.01)	(0.01)	(0.01)
Firm-quarter FE	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes
Region bank FE	Yes	Yes	Yes	Yes
N	440567	440567	440567	440567
R^2	0.853	0.554	0.721	0.915
Adjusted R^2	0.723	0.161	0.476	0.841
F-test statistic	16.332***	3.670***	12.515***	2.042**
degrees of freedom	(13, 37)	(13, 37)	(13, 37)	(13, 37)

Table 1.9: Lagged dependent variable

The table reports robustness tests for the 2SLS model and the reduced form model by including the dependent variable (a dummy equal to one if the loan is nonperforming and zero otherwise) lagged of three quarters among controls. The dependent variable is displayed at the bottom of each column. The independent variables are described in Table A1.1 in Appendix B. For each independent variable the first row reports the coefficient, the second row reports in parenthesis the robust standard error that is corrected for multiclustering at the year-quarter and bank level. Fixed effects are included, "Yes". ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. • denotes rescaled coefficients that have been multiplied by 10^2, whereas •• denotes rescaled coefficients that have been multiplied by 10^5. At the bottom of the table we report the results of standard underidentification and weak identification tests, as well as the Anderson-Rubin test.

	2SLS model		Reduced form model
	First stage	Second stage	
	(1)	(2)	(3)
	Monitor$_{t-2}$	NPL dummy$_t$	NPL dummy$_t$
Monitor$_{t-2}$		-6.826*	
		(3.87)	
IRAP$_{t-6}$	-0.412**		3.758***
	(0.19)		(1.32)
NPL dummy$_{t-3}$	-0.000	0.388***	0.398***
	(0.00)	(0.01)	(0.01)
Share guarantee$_{t-3}$	-0.001*••	0.000	0.000
	(0.00)	(0.00)	(0.00)
Share exposure$_{t-3}$	-0.000	0.016***	0.017***
	(0.00)	(0.00)	(0.00)
Length relation$_{t-3}$	-0.010***•	-0.000	0.036***•
	(0.00)	(0.00)	(0.00)
Capital ratio bank$_{t-6}$	0.019	0.278*	0.187**
	(0.02)	(0.15)	(0.07)
ROA bank$_{t-6}$	-0.001	-0.203	-0.206
	(0.02)	(0.27)	(0.17)
NPL ratio bank$_{t-6}$	-0.005	-0.003	0.003
	(0.02)	(0.14)	(0.08)
Size bank$_{t-6}$	-0.001	0.002	0.009
	(0.00)	(0.01)	(0.01)
Liquidity ratio bank$_{t-6}$	0.008	0.455	0.535**
	(0.04)	(0.31)	(0.24)
Nonretail deposit ratio bank$_{t-6}$	-0.004	-0.002	0.036*
	(0.00)	(0.03)	(0.02)
GDP growth region bank$_{t-6}$	-0.048	-0.362	-0.056
	(0.03)	(0.23)	(0.15)
Employment region bank$_{t-6}$	0.001*	0.005	-0.000
	(0.00)	(0.00)	(0.00)
Inflation region bank$_{t-6}$	-0.001	0.008	0.025***
	(0.00)	(0.01)	(0.01)

Continued on next page

Table 1.9 – *Continued from previous page*

Firm-quarter FE	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes
Region bank FE	Yes	Yes	Yes
N	556717	556717	440567
R^2	0.469		0.873
Adjusted R^2	0.005		0.761
F-test statistic	4.41***	162.901***	155.046***
degrees of freedom	(14, 37)	(14, 37)	(14, 37)
Underidentification test			
Kleibergen-Paap LM statistic	3.99		
Chi-sq P-val	0.046		
Weak identification test			
Kleibergen-Paap Wald F statistic	4.92		
Cragg-Donald Wald F statistic	18.04		
Stock-Yogo critical value 10% maximal IV size	16.38		
Anderson-Rubin test			
Anderson-Rubin Wald statistic		5.78**	
Chi-sq P-val		0.016	

Table 1.10: Loan restructuring

The table reports robustness tests for the 2SLS model and the reduced form model by excluding observations pertaining to loans that are nonperforming at time $t-2$. The dependent variable is displayed at the bottom of each column. The independent variables are described in Table A1.1 in Appendix B. For each independent variable the first row reports the coefficient, the second row reports in parenthesis the robust standard error that is corrected for multiclustering at the year-quarter and bank level. Fixed effects are included, "Yes". ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. • denotes rescaled coefficients that have been multiplied by 10.

	2SLS model		Reduced form model
	First stage	Second stage	
	(1)	(2)	(3)
	Monitor$_{t-2}$	NPL dummy$_t$	NPL dummy$_t$
Monitor$_{t-2}$		-2.651*	
		(1.36)	
IRAP$_{t-6}$	-0.495**		1.417**
	(0.22)		(0.66)
Share guarantee$_{t-3}$	-0.000	0.000	-0.000
	(0.00)	(0.00)	(0.00)
Share exposure$_{t-3}$	-0.000	0.012***	0.012***
	(0.00)	(0.00)	(0.00)
Length relation$_{t-3}$	-0.001***•	-0.000	0.001***•
	(0.00)	(0.00)	(0.00)
Capital ratio bank$_{t-6}$	0.022	0.201**	0.194***
	(0.02)	(0.08)	(0.05)
ROA bank$_{t-6}$	-0.011	-0.115	-0.067
	(0.03)	(0.15)	(0.10)
NPL ratio bank$_{t-6}$	-0.002	-0.046	-0.060
	(0.02)	(0.08)	(0.07)
Size bank$_{t-6}$	-0.001	0.003	0.008**
	(0.00)	(0.00)	(0.00)
Liquidity ratio bank$_{t-6}$	-0.013	0.131	0.323*
	(0.04)	(0.17)	(0.17)
Nonretail deposit ratio bank$_{t-6}$	-0.004*	-0.015	0.010
	(0.00)	(0.02)	(0.01)
GDP growth region bank$_{t-6}$	-0.055	-0.289**	-0.202*
	(0.04)	(0.14)	(0.11)
Employment region bank$_{t-6}$	0.001*	0.005**	0.001
	(0.00)	(0.00)	(0.00)
Inflation region bank$_{t-6}$	-0.001	0.003	0.012
	(0.00)	(0.01)	(0.01)
Firm-quarter FE	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes
Region bank FE	Yes	Yes	Yes
N	490517	490517	384510

Continued on next page

Table 1.10 – *Continued from previous page*

R^2	0.467		0.657
Adjusted R^2	-0.003		0.353
F-test statistic	4.31***	12.466***	15.289***
degrees of freedom	(13, 37)	(13, 37)	(13, 37)
Underidentification test			
Kleibergen-Paap LM statistic	4.09		
Chi-sq P-val	0.043		
Weak identification test			
Kleibergen-Paap Wald F statistic	5.02		
Cragg-Donald Wald F statistic	22.22		
Stock-Yogo critical value 10% maximal IV size	16.38		
Anderson-Rubin test			
Anderson-Rubin Wald statistic		5.77**	
Chi-sq P-val		0.016	

Table 1.11: Different lags

The table reports robustness tests for the 2SLS model and the reduced form model by including the independent variables with different lags. The dependent variable is displayed at the bottom of each column. The independent variables are described in Table A1.1 in Appendix B. For each independent variable the first row reports the coefficient, the second row reports in parenthesis the robust standard error that is corrected for multiclustering at the year-quarter and bank level. Fixed effects are included, "Yes". ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively. • denotes rescaled coefficients that have been multiplied by 10..

	2SLS model			Reduced form model	
	First stage	Second stage			
	(1)	(2)	(3)		(4)
	$Monitor_{t-2}$	NPL $dummy_t$	NPL $dummy_t$		NPL $dummy_t$
$Monitor_{t-2}$		-9.214*			
		(5.17)			
$IRAP_{t-6}$	-0.412**		4.746**	$IRAP_{t-5}$	4.721**
	(0.18)		(1.78)		(1.90)
Share $guarantee_{t-2}$	0.000	0.001	-0.000	Share $guarantee_{t-2}$	-0.000
	(0.00)	(0.00)	(0.00)		(0.00)
Share $exposure_{t-2}$	-0.000	0.026***	0.028***	Share $exposure_{t-2}$	0.028***
	(0.00)	(0.00)	(0.00)		(0.00)
Length $relation_{t-2}$	-0.001***•	-0.000	0.001***	Length $relation_{t-2}$	0.001***
	(0.00)	(0.00)	(0.00)		(0.00)
Capital ratio $bank_{t-6}$	0.019	0.341	0.213**	Capital ratio $bank_{t-5}$	0.183**
	(0.02)	(0.20)	(0.09)		(0.09)
ROA $bank_{t-6}$	-0.003	-0.366	-0.304	ROA $bank_{t-5}$	-0.380**
	(0.02)	(0.33)	(0.20)		(0.19)
NPL ratio $bank_{t-6}$	-0.006	-0.091	-0.064	NPL ratio $bank_{t-5}$	-0.044
	(0.02)	(0.18)	(0.10)		(0.10)
Size $bank_{t-6}$	-0.001	0.003	0.014*	Size $bank_{t-5}$	0.013*
	(0.00)	(0.01)	(0.01)		(0.01)
Liquidity ratio $bank_{t-6}$	0.008	0.509	0.578**	Liquidity ratio $bank_{t-5}$	0.515**
	(0.04)	(0.39)	(0.26)		(0.24)
Nonretail deposit ratio $bank_{t-6}$	-0.004	-0.000	0.047**	Nonretail deposit ratio $bank_{t-5}$	0.046**
	(0.00)	(0.03)	(0.02)		(0.02)
GDP growth region $bank_{t-6}$	-0.048	-0.513*	-0.071	GDP growth region $bank_{t-5}$	-0.167
	(0.03)	(0.28)	(0.17)		(0.17)
Employment region $bank_{t-6}$	0.001*	0.005	-0.001	Employment region $bank_{t-5}$	-0.002
	(0.00)	(0.01)	(0.00)		(0.00)
Inflation region $bank_{t-6}$	-0.001	0.003	0.023**	Inflation region $bank_{t-5}$	0.021**
	(0.00)	(0.01)	(0.01)		(0.01)
Firm-quarter FE	Yes	Yes	Yes		Yes
Bank FE	Yes	Yes	Yes		Yes
Region bank FE	Yes	Yes	Yes		Yes
N	556577	556577	440548		444825
R2	0.470		0.853		0.852
Adjusted R2	0.006		0.724		0.723
F-test statistic	4.54***	8.344***	15.464***		15.849***
degrees of freedom	(13, 37)	(13, 37)	(13, 37)		(13, 38)

	(1)	(2)
Underidentification test		
Kleibergen-Paap LM statistic	4.03	
Chi-sq P-val	0.045	
Weak identification test		
Kleibergen-Paap Wald F statistic	4.98	
Cragg-Donald Wald F statistic	17.96	
Stock-Yogo critical value 10% maximal IV size	16.38	
Anderson-Rubin test		
Anderson-Rubin Wald statistic		6.25**
Chi-sq P-val		0.012

1.7 Appendix A

Optimal lending rate, optimal capital structure, and optimal monitoring effort

The optimal lending rate, r_L^*, the optimal capital structure, k^*, and the optimal monitoring effort, q^*, are obtained by maximizing the bank's expected profits:

$$\underbrace{\max}_{r_L,k,q,0<q\leq 1} \Pi = \left\{ q\left[\left(r_L - r_D\left(1-k\right)\right)\left(1-\tau\right) - r_E k\right] - \frac{1}{2}cq^2 \right\} L\left(r_L\right) \tag{1.7}$$

If the borrowers repay their loans, bank shareholders get a net return after tax of $(r_L - r_D\,(1$ $-k))\,(1-\tau)$. Whereas, if the borrowers do not repay, the bank defaults. In this case, bank shareholders receive nothing and depositors are not repaid because of limited liability. The term $r_E k$ represents the opportunity cost for bank shareholders of investing in the bank, adjusted for the probability of loan repayment.

To solve the model, we consider a sequential process. In the first stage, the lending rate is set so that the bank makes zero expected profits, which is the equilibrium condition of a perfectly competitive market. In the second stage, the bank chooses the optimal level of capitalization. Eventually, in the third stage, the bank determines the desired monitoring effort. We solve the model by backward induction, starting from the last stage. The bank's expected profits can be rewritten as:

$$\underbrace{\max}_{r_L,k,q,0<q\leq 1} \Pi = \left\{ q\left[\left(r_L - r_D\left(1-k\right)\right)\left(1-\tau\right)\right] - \left(r+\xi\right)k - \frac{1}{2}cq^2 \right\} L\left(r_L\right) \tag{1.8}$$

Taking the first order condition with respect to q yields

$$q^* = \frac{\left(r_L - r_D\left(1-k\right)\right)\left(1-\tau\right)}{c}, \quad 0 < q^* \leq 1 \tag{1.9}$$

Equation 1.9 shows one channel through which the corporate tax rate affects the desired level of monitoring. Specifically, an increase in the tax rate reduces bank monitoring via its negative impact on the net return from lending. Note that this effect is partially softened by the fact that a rise in the corporate tax rate entails a decrease in the interest burden of deposits as a higher fraction of these interests are tax deductible. Also, note that a higher lending rate and a higher level of capitalization are associated with higher monitoring incentives. Since the lending rate and the level of capitalization are both endogenous in our model, we need to determine how they are affected by the tax rate before identifying the ultimate effect of the corporate tax rate on bank monitoring.

Under the assumption that depositors are risk-neutral, they require a return equal to $r_D = \frac{r}{\mathbb{E}[q|k]}$. The denominator represents depositors' expectations about bank monitoring (and, hence, the survival probability of the bank), as inferred by the level of capitalization. As in Dell'Ariccia et al. (2014), we assume that these expectations are correct in equilibrium, meaning that $\mathbb{E}\left[q\,|k\right] = q^*$. Substituting r_D in equation 1.9, we solve for the optimal monitoring effort:[35]

$$q^* = \min\left\{\frac{r_L\left(1-\tau\right)+\sqrt{r_L^2\left(1-\tau\right)^2-4rc\left(1-k\right)\left(1-\tau\right)}}{2c}, 1\right\} \tag{1.10}$$

We can now maximize bank expected profits with respect to the level of capitalization, subject to the equilibrium condition of depositors $r_D = \frac{r}{q^*}$

$$\underbrace{\max}_{k}\Pi = \left[\left\{q^* r_L\left(1-\tau\right) - r\left(1-\tau\right) - k\left(r\tau+\xi\right) - \frac{1}{2}cq^{*2}\right]L\left(r_L\right) \tag{1.11}$$

This yields

$$\begin{aligned}\frac{\partial \Pi}{\partial k} &= \frac{\partial q^*}{\partial k} r_L\left(1-\tau\right) - r\tau - \xi - \frac{\partial q^*}{\partial k} cq^* = \\ &= \frac{r_L r\left(1-\tau\right)^2}{2\sqrt{r_L^2\left(1-\tau\right)^2 - 4rc\left(1-k\right)\left(1-\tau\right)}} - r\tau - \xi - \frac{r\left(1-\tau\right)}{2} \overset{!}{=} 0\end{aligned} \tag{1.12}$$

We then solve for k obtaining

$$k^* = 1 - r_L^2\left(1-\tau\right)\frac{\left(\xi+r\right)\left(\xi+r\tau\right)}{rc\left(r\tau+2\xi+r\right)^2} \tag{1.13}$$

We now derive the optimal lending rate imposing the zero profit condition, stemming from the assumption of a perfect competitive market

$$\left[q^* r_L\left(1-\tau\right) - r\left(1-\tau\right) - k^*\left(r\tau+\xi\right) - \frac{1}{2}cq^{*2}\right]L(r_L) \overset{!}{=} 0 \tag{1.14}$$

[35] The optimal level of monitoring is obtained by solving a quadratic equation. Formula 1.10 consists in the root having the highest value. We select this root as, keeping everything else equal, both the bank and the borrowers are better off with higher monitoring.

To solve equation 1.14 we have to substitute the expressions for q^* and k^*. First, we replace k^* in the expression for q^* so that

$$q^* = \frac{r_L (1-\tau)(r+\xi)}{c(r\tau + 2\xi + r)} \tag{1.15}$$

Then, we substitute the resulting q^* and k^* in equation 1.14 and we solve for the optimal lending rate

$$r_L^* = \sqrt{\frac{2cr(r\tau + 2\xi + r)^2}{(1-\tau)(3r\xi + r^2 + 2\xi^2 + r^2\tau + r\tau\xi)}} \tag{1.16}$$

From that we can directly obtain the derivative of the desired lending rate with respect to the corporate tax rate

$$\frac{\partial r_L^*}{\partial \tau} = \sqrt{\frac{2cr(r^2+\xi^2)^2 (r + 2\xi + r\tau)^2}{(1-\tau)^3 (3r\xi + r^2 + 2\xi^2 + r^2\tau + r\tau\xi)^3}} > 0 \tag{1.17}$$

This derivative is positive, suggesting that, when the corporate tax rate increases, the bank shifts part of the tax burden on its borrowers by rising the lending rate. Then, substituting r_L^* in the expression for k^*, we obtain the optimal level of capitalization

$$k^* = \frac{(r^2 + r\xi)(1-\tau)}{3r\xi + r^2 + 2\xi^2 + r^2\tau + r\tau\xi} \tag{1.18}$$

Its derivative with respect to the corporate tax rate is

$$\frac{\partial k^*}{\partial \tau} = -\frac{(r^2 + r\xi)(4r\xi + 2r^2 + 2\xi^2)}{(3r\xi + r^2 + 2\xi^2 + r^2\tau + r\tau\xi)^2} < 0 \tag{1.19}$$

The negative sign implies that an increase in the corporate tax rate reduces the level of capitalization.

Finally, we can retrieve the optimal monitoring effort by substituting r_L^* in equation 1.15 and calculate its derivative with respect to the corporate tax rate. This yields

$$q^* = \sqrt{\frac{2r(r+\xi)^2(1-\tau)}{c(3r\xi + r^2 + 2\xi^2 + r^2\tau + r\tau\xi)}} \tag{1.20}$$

$$\frac{\partial q^*}{\partial \tau} = -2(r+2\xi)(r+\xi)\sqrt{\frac{r^3}{c(1-\tau)(3r\xi + r^2 + 2\xi^2 + r^2\tau + r\tau\xi)^3}} < 0 \tag{1.21}$$

1.8 Appendix B

Table A1.1: Description of variables used in the empirical analysis

Variable name	Description
Loan-level variables	
Past-due dummy	A dummy variable equal to one if the loan is past-due by 90 days or more, and zero otherwise.
UTP dummy	A dummy variable equal to one if the loan is defined as "unlikely-to-pay", UTP, meaning that the bank envisages the possibility that the loan will not be repaid in full.
Bad loan dummy	A dummy variable equal to one if the loan is defined as "bad loan", meaning that the bank considers the loan as impaired
NPL dummy	A dummy variable equal to one if the loan is defined either as past-due by 90 days or more, "unlikely-to-pay", or "bad loan".
Monitor	Total number of requests for information made by a bank on an existing borrower in the quarter. To build this variable we aggregate all requests for information without making any distinction regarding the reason.
Share guarantee	Value of the credit guarantee (collateral or personal guarantee) as a fraction of the loan.
Share exposure	The ratio between the amount of the loan extended to the firm by the individual bank and the total borrowing of the firm.
Length relation	Duration of the credit relationship expressed in quarters.
Firm-level variables	
Size firm	Firm size, computed as the logarithm of total assets.
ROA firm	Firm profitability, calculated as the ratio of net income to total assets.
Credit score firm	The credit score assigned to the firm by the provider of the CERVED database. Credit score values range from 1 to 9. A credit score of 1 corresponds to firms with the highest credit quality, while a credit score of 9 corresponds to firms which are essentially in default.
Capital ratio firm	The equity-to-asset ratio of the firm.

Continued on next page

Table A1.1 – *Continued from previous page*

Variable name	Description
Industry firm	The “ateco” code identifying the industry of the firm. Industry takes 5 different values, each one corresponding to a specific sector.
Bank-level variables	
Size bank	Bank size, computed as the logarithm of total assets.
Capital ratio bank	The equity-to-asset ratio of the bank.
Liquidity ratio bank	The ratio of liquid assets to total assets of the bank
ROA bank	Bank profitability, calculated as the ratio of net income to total assets.
Nonretail deposit ratio bank	The ratio of nonretail deposits to total deposits of the bank.
NPL ratio bank	The fraction of nonperforming loans to total loans to the private sector.
Regional variables	
IRAP	The regional IRAP tax rate applied to the bank.
GDP growth region firm	GDP growth of the firm’s region, computed as the first difference in the logarithm of GDP. GDP is deflated using CPI with 2010 as the reference year.
Employment region firm	Employment rate of the firm’s region, expressed in percentage points.
Inflation region firm	Inflation rate of the firm’s region, expressed in percentage points.
GDP growth region bank	GDP growth of the bank’s region, computed as the first difference in the logarithm of GDP. GDP is deflated using CPI with 2010 as the reference year.
Employment region bank	Employment rate of the bank’s region, expressed in percentage points.
Inflation region bank	Inflation rate of the bank’s region, expressed in percentage points.

Chapter 2

Fixed Rate versus Adjustable Rate Mortgages: Evidence from Euro Area Banks

Why do residential mortgages carry a fixed or an adjustable interest rate? To answer this question we study unique data from 103 banks belonging to 73 different banking groups across twelve countries in the euro area. To explain the large cross-country and time variations observed, we distinguish between the conditions that determine the local demand for credit and the characteristics of banks that supply credit. As bank funding mostly occurs at the group level, we disentangle these two sets of factors by comparing the outcomes observed for the same banking group across the different countries. Local demand conditions dominate. In particular we find that the share of new loans with a fixed rate is larger when: (1) the historical volatility of inflation is lower, (2) the correlation between unemployment and the short-term interest rate is higher, (3) households' financial literacy is lower, and (4) the use of local mortgages to back covered bonds and mortgage-backed securities is more widespread.

Albertazzi, U., Fringuellotti, F., Ongena, S., 2018. Fixed Rate versus Adjustable Rate Mortgages: Evidence from Euro Area Banks. Working Paper.

2.1 Introduction[1]

Conventional mortgages can be classified in two main types: fixed rate mortgages and adjustable rate mortgages. Fixed rate mortgages (FRMs) charge a nominal interest rate that does not change during the entire life of the loan. Adjustable rate mortgages (ARMs) charge an interest rate that is tied to a benchmark and varies over time. Households that select an ARM are exposed to the short-term variability in the periodic payments required by this type of mortgage (Campbell and Cocco, 2003). The volume of FRMs and ARMs extended to households in the economy depends on a broad set of factors that affect the demand of borrowers and the supply of lenders (ECB, 2009).

A striking feature of the credit market in the euro area is the very large heterogeneity across countries in the granting of fixed versus adjustable rate mortgages. FRMs are dominant in Belgium, France, Germany and the Netherlands, while ARMs are prevailing in Austria, Greece, Italy, Portugal and Spain (ECB, 2009; Campbell, 2012). The variation in the share of FRMs to total new mortgages differs across countries as well, with little variation over time in Germany and Portugal, but far more in Italy and Greece (ECB, 2009). This observed variation across countries and over time has three major implications. First, the transmission of monetary policy is heterogeneous across countries. Being a major liability in the balance sheet of most households, mortgages likely play a key role in the transmission of monetary policy (Di Maggio et al., 2017). This is especially true in systems where ARMs are dominant because, on top of the traditional bank lending channel, also the floating rate channel is at work, with significant macroeconomic effects.[2] Second, the allocation of interest-rate risk between the banking sector and the real sector differs across countries, with direct consequences for financial stability. Third, the effectiveness of macroprudential policies in containing mortgage defaults varies across countries, with potential repercussions for the financial system and the real economy (Stanga et al., 2017).[3] In light

[1]The opinions expressed in this paper are those of the authors only and do not necessarily reflect those of the Bank of Italy and the ECB.

[2]Ippolito et al. (2017) define the floating rate channel as the mechanism whereby conventional monetary policy actions are transmitted directly to borrowers' balance sheet via a change in the interest rate paid on outstanding (indexed) loans.

[3]Stanga et al. (2017) show that restrictive macroprudential policies are negatively associated with mortgage delinquencies in countries where FRM are prevalent.

of that, investigating the determinants of the prevalent type of mortgage across countries and over time is crucial in order to derive normative insights.

In this paper we exploit unique bank-level information on lending activity in the euro area in order to understand what drives the prevalence of FRMs or ARMs. In particular, we investigate to what degree the wide cross-country heterogeneity in the prevalent interest rate type of mortgage is driven by differences in demand or supply conditions.

From a methodological perspective, we distinguish the role played by borrower specific characteristics from that of bank specific factors. The former include all features that make borrowers demand one or the other type of mortgage, as well as those that make a borrower more or less suitable to be financed at a fixed rate.[4] The latter include funding and liquidity conditions, which may influence the ability of banks to supply FRMs.[5]

Our identification strategy relies on the assumption that funding takes place at the consolidated level. Lending policies are mainly influenced by fund-raising and liquidity conditions. Funding is defined and mainly occurs at the consolidated level.[6] Thus, the ability and willingness of a banking group to grant loans with certain features is also mainly determined at the consolidated level, particularly when the group operates in a monetary union, such as the euro area. Similar considerations apply to bank liquidity. Our assumption

[4] The riskiness of the lending exposure determines whether a mortgage can be financed through long-term funds at a fixed rate, for example, by issuing covered bonds or mortgage-backed securities. If a loan can be used to back covered bonds or mortgage-backed securities, the bank can offer a more convenient fixed interest rate.

[5] Typically, banks rely on short-term funding at adjustable rate. A natural consequence is that banks are more willing to supply ARMs. But to the extent that they can raise long-term funds at fixed rate, banks are also able to supply FRMs. This holds true as long as banks keep an exposure to interest rate risk, as documented by Hoffmann et al. (2017). Indeed, if banks were to fully hedge, they would be equally willing to supply FRMs and ARMs. Analysing bank specific characteristics allows us also to shed light on banks' exposure to interest rate risk.

[6] Cross-border banks define their funding mix as to minimize the cost of capital (Gu et al., 2015). Long-term funding instruments are issued taking into account differences across countries in terms of taxation, regulation, quality of required services and infrastructures, as well as development of capital markets. For example, banks can raise funds through cross-border securitisation or concentrating covered bonds issuance in certain countries. Despite cross-border banks can select different funding models, funding mainly occurs at the consolidated level. While showing a shift towards a more decentralized funding at the onset of the recent financial crisis, Gambacorta et al. (2017) document that cross-border banks' liabilities from foreign affiliates amount only to 41% of total funds raised overseas.

is in line with the focus of market investors and regulators on consolidated bank balance sheets and with the literature on cross-border banks as shock propagators, where local lending conditions are affected by shocks to the consolidated balance sheet (Cetorelli and Goldberg, 2011; Schnabl, 2012; Célérier et al., 2018).[7]

Intuitively, this allows us to disentangle borrower demand from bank supply by comparing, on the one hand, the lending patterns observed for the same cross-border banking group in different economies and, on the other hand, the lending patterns observed across different cross-border banking groups operating in the same economy.

In practice, we decompose the variation of the share of FRMs to total new mortgages, henceforth abridged with "share of FRMs", into "country demand factors" and "bank supply factors", using a fixed effects model and exploiting cross-border banking groups. This approach is close in spirit to Amiti and Weinstein (2016) and Greenstone et al. (2015). Country demand factors capture specific features of the borrowing country which are more related to loan demand, that is to the characteristics of borrowers in that country, whilst bank supply factors capture funding and liquidity conditions which are relevant for lending supply.

One main advantage of our approach is that we are able to jointly investigate the role played by demand and supply conditions. Moreover, we are not bound to select a list of proxies for demand and supply factors, as typically done in the literature. Making such a selection is difficult as one cannot be sure that the list is exhaustive and, more importantly, that the variables under consideration truly capture only demand or only supply.[8] On the negative side, our estimated country demand factors are not directly interpretable in economic terms, as they are likely to encompass a heterogeneous set of variables. Thus,

[7] Bank supervision activity almost exclusively focuses on consolidated balance-sheet conditions, including the level of interest rate risk (BCBS, 2012; ECB, 2014). Additionally, the design of banks' surveys is typically aimed at gauging lending standards at the consolidated level.

[8] There exist factors that in principle may exert a role in shaping both demand and supply conditions of FRMs, relative to ARMs. This is the case, for example, for legislation on issuance of covered bonds. Namely, if its effect is to allow banks to issue such instruments, then it is exerting an effect on the supply of FRMs. If instead its effect is to make a mortgage issued locally eligible to be used as collateral for covered bonds, for instance due to specific requirements in terms of loan-to-value (ECBC Covered Bond Comparative Database; ECB, 2008; ECBC, 2016), then it is exerting an effect on the demand of FRMs. For these reasons, it is difficult to separate demand from supply based on pre-selected lists of proxies for the two sides of the market.

as a second step, and in the same vein of Ongena and Smith (2000), we adopt a two-stage approach whereby the estimated demand factors are regressed on variables that are economically motivated.

Our main finding indicates a prominent role for country demand factors which explain almost 72% of the total variation in the share of FRMs observed in the sample, as opposed to 19% associated with bank supply factors (the remaining 9% being the variation that the model is unable to explain). A number of robustness exercises show that this result is confirmed when we use a larger dataset including smaller and domestic institutions, as well as when we adopt a non-linear model specification.

In a first extension of the baseline regressions we explore more in detail the time variation in the share of FRMs, which turns out to be strongly and negatively correlated with the term spread, that is the slope of the yield curve. In line with the main findings, the results of this exercise suggest that changes in the term spread mainly entail changes in the demand for FRMs, relatively to ARMs. Specifically, 79% of the variation in the share of FRMs driven by the term spread is ascribable to demand factors. The elasticity of demand on the term spread differs across countries.

We more broadly explore the economic variables behind the cross-country differences in local demand conditions, according to the two-stage procedure, as described above. The variables selected are taken from the existing literature, but we also put emphasis on a novel variable that has not been considered so far. We start from the observation that if households expect to be unemployed when interest rates are low, the ARM provides households with an insurance coverage (while the FRM does not). This simple (but at first sight somewhat counterintuitive) remark leads us to check whether the share of FRMs is related to the correlation between the unemployment rate and the short-term interest rate. This correlation turns out to be highly significant and economically relevant in explaining the demand component of the share of FRMs. Specifically, an increase in the correlation between the unemployment rate and and the short-term interest rate by one standard deviation (an increase of 0.49) leads to an increase of 14 percentage points in the average share of FRMs per country explained by demand conditions.

Concerning the statistical significance of the other (more standard) economic factors underlying the demand (having controlled for supply side factors), we document a role for financial literacy, whose effect turns out to be negative, in line with the notion that more educated borrowers can better understand complex financial products such as ARMs.[9] One standard deviation increase in financial literacy (an increase of 8 percentage points) entails a decrease of 42 percentage points in the average share of FRMs per country. Households in countries where the covered bonds market is more developed are more likely to borrow at a fixed rate, given that such bank funding instruments backed by mortgages are typically issued at long maturities and at fixed rates.[10] For a similar reason, also the volume of securitized mortgages entails a higher likelihood of households selecting a FRM.[11] An increase in the outstanding amount of mortgage covered bonds and residential mortgage-backed securities, scaled by GDP, by one standard deviation (corresponding to 6 percentage points for both) leads to an increase of 32 and 17 percentage points, respectively, in the average share of FRMs per country explained by the demand. Finally, high historical volatility of inflation is strongly and negatively related to the share of FRMs, consistently with the idea that the macroeconomic history of a country affects households' mortgage choice.[12] A one standard deviation increase in the historical inflation volatility (an increase of 9 percentage

[9] Financially educated borrowers are more familiar with the concepts of fixed interest rate, adjustable interest rate and interest compounding. As such, they are able to grasp that the interest rate applied on an ARM and that of a FRM are not equivalent at the inception of the loan. Indeed, the interest rate on a FRM embeds not only the expectation of the future short-term interest rate, but also a term premium and the cost of the prepayment option (Campbell and Cocco, 2003). Selecting an ARM rather than a FRM allows to avoid these add-ons.

[10] Funding via covered bonds is a factor that could clearly indicate both shifts in demand (i.e., borrower specific) and shifts in supply (i.e., lender specific). Although the supply side might play a stronger role, what are we capturing in our setting is whether households' characteristics in a given country make mortgages more ore less eligible to secure covered bonds. As such, we are focusing on the demand component of this factor.

[11] Banks engagement in loan securitization can be driven both by demand and supply conditions. Since we control for the supply side, our factor catches only the demand component that is of major interest.

[12] Countries with higher volatility of inflation before the introduction of the euro were characterized by a strong prevalence of ARMs. This is in line with the idea that, if a FRM can be prepaid without penalties, high inflation risk leads banks to reduce the supply of FRMs by increasing the interest rate applied on such loans. As a consequence households are more likely to select ARMs (Campbell, 2012; Badarinza et al., 2017). Alternatively, this may signal the existence of a stronger insurance motive attached to ARMs (countries with higher inflation risk are those where households are more likely to be unemployed when the short-term interest rate is very low). The fact that ARMs continue to dominate the mortgage market of these countries even after the entry to the eurozone suggests a certain stickiness in households' behavior (Campbell, 2012).

points) entails a decrease of 59 percentage points in the average share of FRMs per country.

We complete our study adopting a similar approach to explain prices instead of quantities, that is considering as dependent variable the spread between FRMs and ARMs interest rates, rather than the share of FRMs. Our findings indicate that also the spread between FRMs and ARMs interest rates is mainly driven by demand conditions.

The reminder of the paper is organized as follows. Next section reviews the existing literature and explains the contribution of this work. Section 2.3 discusses the identification strategy. Section 2.4 describes the dataset. Section 2.5 presents the methodology and the results of the analysis on the share of FRMs. Section 2.6 integrates the preceding with some robustness checks. Section 2.7 presents the results of the analysis on the spread between FRMs and ARMs interest rates. Section 2.8 concludes.

2.2 Literature and Contribution

2.2.1 Demand and Supply Factors

The existing literature provides both theoretical modeling and empirical evidence on the determinants of the prevalent type of mortgage. A wide range of demand factors and supply factors may drive the choice between FRMs and ARMs.

As for demand factors, an important role is ascribed to borrower's financial condition and level of education. In a pioneering work, Campbell and Cocco (2003) derive relevant theoretical predictions by treating mortgage choice as a problem in household risk management. In their framework, households subject to binding borrowing constraints at the time of the loan application, such as low income and low level of savings, are likely to choose the loan with the lowest interest rate. In general, this is then an adjustable rate as a fixed interest rate will include a term premium and the cost of the prepayment option.[13] Yet, an ARM exposes households to the income risk of short-term variability in the periodic payments. Thus, households with a limited income risk bearing capacity, for example in

[13] The interest rate on an ARM is close to the short-term interest rate. The interest rate on a FRM is related, instead, to the long-term interest rate. The existence of a term premium and a cost of early repayment means that the interest rate on a FRM is not equivalent to the expectation of the future short-term interest rate. As a consequence, at inception of a loan the interest rate on an ARM and the interest rate on a FRM are not equivalent.

case of high loan-to-income ratio and low financial wealth, are likely to select a FRM.

Several empirical papers have extensively investigated the role of income, savings, indebtedness and financial wealth in the choice of housing loans relying on households' income and wealth surveys (Paiella and Pozzolo, 2007; Fornero et al., 2011; Ehrmann and Ziegelmeyer, 2017). These studies provide a general support for the predictions of Campbell and Cocco (2003).

Borrowers' education, especially the degree of financial literacy, is an important driver of mortgage choice as well (Agarwal et al., 2010; Fornero et al., 2011; Gathergood and Weber, 2017). In general, more educated borrowers have a deeper understanding of the intrinsic features of ARMs and FRMs. On the one hand, they are aware that, unconditionally, a FRM is more expensive than an ARM, due to the term premium and the cost of the prepayment option mentioned above. For this reason, they are more likely to select an ARM (Agarwal et al., 2010; Gathergood and Weber, 2017). On the other hand, they are also mindful of the potential exposure to income risk if they choose an ARM (Fornero et al., 2011).

Supply factors consist in bank funding and liquidity conditions. In general, the composition of liabilities affects, and is affected, by the type of loan a bank is more willing to offer and thus the quoted interest rates (Kirti, 2017). A few empirical studies indeed show that lower bank bond spreads, lower deposit pass-through, lower exposure to interest rate risk and higher access to securitization make banks more prone to extend fixed rate loans (Fuster and Vickery, 2014; Foà et al., 2015; Basten et al., 2017).

Beside these rather intuitive factors, there exist a set of macroeconomic factors that exert their effects either through demand or supply. These include current and future expected interest rates, as well as the unemployment rate and the macroeconomic history of a country.

The current spread between the interest rates on FRMs and ARMs is a leading factor of mortgage choice (Paiella and Pozzolo, 2007; Koijen et al., 2009; Fornero et al., 2011; Badarinza et al., 2017). This suggests that households behave myopically, selecting the type of loan that requires the lowest payments at the time of the loan application. However, households' expectations on the future interest rate applied on ARMs play a role as well, but only over the short horizon of one year (Koijen et al., 2009; Foà et al., 2015; Badarinza et al., 2017).

The difference between long-term and short-term interest rates is a component of the spread between FRMs and ARMs interest rates. As such, the current term spread is also a determinant of mortgage choice (Koijen et al., 2009; Basten et al., 2017; Ehrmann and Ziegelmeyer, 2017). Since in the literature on the bank lending channel the level of interest rates is recognized to be able to shift both the demand and the supply of credit, one can surmise that the term spread may act as a shifter of both the demand and the supply of FRMs, relatively to ARMs.

The historic volatility of inflation plays an important role in the choice of mortgages as well. Countries with a history of high volatility of inflation prior to the introduction of the euro show a prevalence of ARMs (Campbell, 2012; Badarinza et al., 2017). This persists even after the adoption of the euro, suggesting a substantial inertia in households' behavior (Campbell, 2012).

The volatility of the unemployment rate, as a proxy for households' expected income, is an additional driver of the prevalent type of mortgage. In countries with high volatility of the unemployment rate households are more likely to select a FRM, as future income is expected to be unstable (Ehrmann and Ziegelmeyer, 2017).

Guren et al. (2018) emphasize the prominent role in mortgage choice of the monetary policy reaction function to aggregate shocks. If the central bank decreases interest rates in response to a crisis, an ARM provides households with higher insurance benefits allowing a higher degree of consumption smoothing. We are the first to test empirically this prediction including among our country demand factor a novel variable, namely the correlation between the unemployment rate and the short-term interest rate.

Table 2.1 summarizes all the determinants of mortgage choice identified in the literature, as well as those analysed in this study.

[Insert Table 2.1 here]

2.2.2 Contribution

The existing literature investigates the plethora of factors driving the choice between FRMs and ARMs, mainly focusing on one specific country and without providing information on the relative importance of demand and supply factors. To the best of our knowledge, the works of Ehrmann and Ziegelmeyer (2017) and Badarinza et al. (2017) are the only two papers to examine the determinants of mortgage choice across countries.

Using a new household wealth survey, the Eurosystem household finance and consumption survey, Ehrmann and Ziegelmeyer (2017) provide a deep investigation of the demand side, but ignore completely the supply side. Relying on monthly country-level information, Badarinza et al. (2017) analyse how current and future expected interest rates affect the time variation in the share of ARMs to total new mortgages. They partially investigate the cross-country variation as well, but look exclusively at the role played by the historic volatility of inflation. Both these studies are not able to investigate jointly the broad spectrum of demand and supply factors driving mortgage choice, neither to disentangle them.

We are able to overcome these limitations by using unique granular bank-level information on a sample of intermediaries operating in twelve countries in the euro area. The structure of our dataset allow us to take a step towards identifying demand and supply of FRMs, relatively to ARMs. Specifically, we rely on an identification strategy that utilizes cross-border banking groups to disentangle country demand factors from bank supply factors. In this way we are able to rigorously examine to what extent the wide cross-country heterogeneity and time variation in the prevalent interest rate type of mortgage is driven by demand or supply conditions.

Assessing the relative importance of demand and supply is crucial because the policy implications may differ substantially depending on what is the actual driver. Eventually, we are the first to explore the role of demand and supply also on the relative price of FRMs and ARMs.

2.3 Identification

Our identification strategy builds on the idea that funding takes place at the consolidated level. This allows us to disentangle demand from supply by comparing the lending behavior of the same cross-border banking group in different countries, as well as the lending behavior of different cross-border banking groups operating in the same economy.

Our identification strategy is supported by several facts. First, lending policies are mainly driven by bank funding and liquidity conditions. In a cross-border banking group funding is defined at the consolidated level as to minimize the cost of capital. For example, Gu et al. (2015) show that international banks raise debt through subsidiaries operating in countries with a more favorable tax system. In general, cross-country differences in terms of taxation, regulation, bureaucracy, services and infrastructure, as well as development of capital markets have a crucial role in the way banks issue long-term funding instruments. For instance, international banks can raise funds relying on cross-border securitisation or concentrating covered bonds issuance in certain countries. Indeed, covered bonds legislations in most European countries, with the exception of Greece and the Netherlands, allow to include mortgages originated abroad (typically in the European Economic Area and in Switzerland, or more broadly in OECD countries) in the covered pool (ECBC Covered Bond Comparative Database; ECB, 2008; ECBC, 2016). Additionally, in a cross-border banking group funding mainly occurs at the consolidated level. Although international banks have progressively adopted a more decentralized funding model after the recent financial crisis, Gambacorta et al. (2017) show that cross-border banks' liabilities from foreign branches and subsidiaries represent, even recently, still 41% of total funds raised abroad. For similar reasons, also liquidity conditions are defined at the consolidated level. As a consequence, the ability and willingness of a cross-border banking group to grant loans with given characteristics is also mainly determined at the consolidated level. This is especially true if the group operates in a monetary union, such as the euro area, characterized by homogenous regulations and integrated capital markets.

Second, when looking at cross-border banks, market investors and regulators are mainly focused on consolidated balance sheets. For example, the "core principles for effective banking supervision" depicted by the Basel Committee on Banking Supervision markedly refer to the assessment of consolidated balance sheet conditions, also regarding the exposure to interest rate risk (BCBS, 2012). These principles are broadly confirmed by the ECB guide to banking supervision (ECB, 2014). Additionally, the design of banks' surveys is typically aimed at gauging lending standards at the consolidated level. This is the case, for example, of the Euro Area Bank Lending Survey and the Senior Loan Officer Opinion Survey run by the Eurosystem and by the Federal Reserve System, respectively.

Third, our identification assumption is consistent with the literature on cross-border banks as shock propagators. This literature shows that funding and liquidity shocks to the holding of a cross-border banking group affect local lending supply (Cetorelli and Goldberg, 2011; Schnabl, 2012; Célérier et al., 2018).

While it is reasonable to argue that lending policies are mainly driven by funding and liquidity conditions of the banking group, we cannot exclude that local funding or other factors may affect bank supply at the country level. For example, local subsidiaries may experience a certain degree of flexibility, which would be subsumed in our country demand factors. However, the fact that fund-raising and liquidity conditions are prominent determinants of lending supply, as well as the fact that they are mostly defined at the consolidated level, ensures that our identification strategy is reliable.

More importantly, we cannot exclude that cross-border banks define local lending policies taking into account the demand conditions that are specific to each country in which they operate. For example, it could be the case that a bank is less willing to extend ARMs in an economy characterised by high default rates (if it thinks that granting ARMs would entail even higher default rates). Our methodology is not able to isolate such component of lending supply that varies with borrowers' characteristics; nonetheless, it can effectively identify supply conditions driven by bank funding, sometimes referred to as pure supply factors, which is the objective of our analysis. In this respect, our analysis shares exactly the same advantages and limitations of studies exploiting more granular data to control for credit demand conditions.[14]

2.4 Data

This paper uses the Individual Monetary and Financial Institution Interest Rates (IMIR) dataset held by the Bank of Italy. This dataset includes monthly bank-level information on a representative sample of seventy-three monetary and financial institutions (MFIs),[15] which we will henceforth simply call "banks", acting in twelve countries in the euro area. In particular our panel includes banks operating in Austria, Belgium, France, Germany, Greece, Italy, Latvia, Luxembourg, Portugal, Slovenia, Spain and the Netherlands. Data cover the period that goes from July 2007 to December 2015. The available information encompasses the amount granted and a weighted average of the interest rate applied to

[14]For example, if banks apply tighter lending criteria to small size borrowers, such extra tightening is captured by borrowers-time fixed effects, which are typically meant to control for demand conditions.

[15]According to the European Central Bank monetary and financial institutions are resident credit institutions as defined in European Union law, and other resident financial institutions whose business is to receive deposits and/or close substitutes for deposits from entities other than MFIs and, for their own account (at least in economic terms), to grant credits and/or make investment in securities.

new mortgages. Overall, we have 103 banks associated to 73 banking groups. The latter include five cross-border banking groups. Detailed information on our dataset is exposed in Table 2.2.

[Insert Table 2.2 here]

Figure 2.1 shows the average share of FRMs, the average spread between FRMs and ARMs interest rates, and the term spread computed as the difference between the 10-year Interest Rate Swap rate and the 3-month Overnight Index Swap rate. Looking at the average share of FRMs, we find a substantial cross-country heterogeneity. We can divide countries in two main groups. France, Germany and the Netherlands exhibit a large proportion of FRMs over the entire time period of our analysis. All the other countries exhibit more time variation and for most of them the average share looks negatively related to the average spread. Looking at the spread between FRMs and ARMs interest rates, some differences are observable as well, although for this metric the heterogeneity seems contained. The time patterns of the average spread largely reflect those of the slope of the term structure as measured by the term spread.

[Insert Figure 2.1 & 2.2 here]

Figure 2.2 displays the evolution of the share for domestic and foreign banks within countries, for the two representative group of economies. The heterogeneity across banks within (these groups of) countries is non negligible, but still much smaller than what is observable across such (groups of) countries. In both groups of economies foreign banks behave consistently with the domestic banks of the country in which they operate. This evidence suggests that country factors may play a major role than bank supply factors.

Table 2.3 reports basic statistics for the share of FRMs and the spread between FRMs and ARMs interest rates for each country in our data set.

[Insert Table 2.3 here]

2.5 Empirical Analysis

2.5.1 Baseline Model

Our methodology relies on the approach proposed by Amiti and Weinstein (2016), although applied to our unique dataset, and exploits cross-border banking groups to decompose the share of FRMs into demand and supply components.[16] More specifically, we estimate the following type of regression:

$$share(b,c,t) = \alpha(c,t) + \beta(h(b),t) + \varepsilon(b,c,t) \tag{2.1}$$

In equation 2.1 the share of FRMs extended by a given bank b operating in a given country c at time t is regressed on a set of different fixed effects. The terms $\alpha(c,t)$ represent month-country fixed effects. They consist in all observable and unobservable time varying and time invariant characteristics of country c and, as such, they are meant to capture the demand conditions prevailing in that economy. Obviously, no other country specific controls can be added to the specification, as these would be subsumed in the month-country fixed effects. This means that the inclusion of month-country fixed effects in equation 2.1 is equivalent to the use of an arbitrarily large set of country macroeconomic controls, which is why we argue that we are effectively capturing country demand factors. Nonetheless, their limitation in this context is related to the inability to control for demand conditions that are specific to individual intermediaries. As most of our analysis focuses on cross-border banks, and since these are typically large banks operating on a national scale and with a diversified set of borrowers, we consider our approach appropriate. The terms $\beta(h(b),t)$ represent month-banking group fixed effects, $h(b)$ denoting the holding of bank b. They consist in all observable and unobservable time varying and time invariant characteristics of banking group h and, as such, they are aimed at capturing bank supply conditions. In light of the fact that lending policies are usually defined at the consolidated level taking into account the financing conditions of the entire group, we argue that this set of fixed effects reasonably accounts for bank supply factors.[17]

[16] Greenstone et al. (2015) adopt a similar methodology, but they decompose the variation of their dependent variable using time invariant rather than time varying fixed effects.

[17] Cross-border banks may sort themselves in countries that share similar characteristics. Even within a country, they may specialize in lending to households that demand a certain type of mortgage. If this is the

By construction, equation 2.1 can only be estimated in the subsample of observations pertaining to cross-border banks. In this sample, equation 2.1 provides the upper limit of the R^2 that is achievable by regressing the share of FRMs on any set of variables capturing (time varying) characteristics of the borrowing country c and (time varying) characteristics of the lender h. Ideally, we would control for supply factors at the bank level, as we cannot exclude the possibility that some of these intermediaries experience some degree of autonomy (Houston et al., 1997). We investigate whether this is the case by estimating alternative specifications to model 1 where we can say something about the role of supply factors defined at the individual bank level. Of course this comes at some cost, as it requires to abandon the use of time varying fixed effects. We evaluate the size of costs associated with this approximation. Eventually, in order to exploit the information available in the entire sample, we also explore simpler specifications where the set of controls is less fine that what is implied in model 2.1.

2.5.2 Baseline Results

Models 1-3 of Table 2.4 report three specifications in which the share of FRMs is regressed on, respectively, month-country fixed effects, month-banking group fixed effects and both of these sets of fixed effects jointly. The latter is exactly the model specified in equation 2.1. Month-country fixed effects explain a significant fraction of the variation in the share (84%), suggesting a prominent role of demand factors. When considered alone, month-banking group fixed effects also explain some of the variation in the dependent variable (32%), but significantly less than month-country fixed effects. If taken together these two sets of fixed effects can explain 91% of total variation in the share. By decomposing the R^2 of model 3 according to the Shorrocks-Shapely approach, we find that the component of R^2 related to month-country fixed effects (72%) is considerably higher than the component related to month-banking group fixed effects (19%), confirming that demand conditions play a prominent role.[18] When saturating the previous specification by including also bank (time

case, our banking-group fixed effects may capture demand rather than supply factors. Nevertheless, the set of cross-border banks that we exploit in our regression analysis includes big universal banks which operate in countries that show a significant difference in the prevalent type of mortgage. Such big players are likely to operate on a national scale without specializing in a specific type of mortgage.

[18]In the fixed effect decomposition of model 3 we have 360 month-country dummies versus 393 month-banking group dummies. The two sets of fixed effects are well balanced, meaning that the results are

invariant) fixed effects, as in model 4, we are able to explain almost the entire variation in the dependent variable. Even if we interpret these dummies as (time invariant) supply factors at the bank level, we would still conclude that overall supply conditions explain only a minor portion of the total variation in the share of FRMs.

[Insert Table 2.4 here]

One may be concerned whether the specific sample over which we are able to conduct our exercise, which is given by all observations (bank-month pairs) pertaining to cross-border banking groups, is representative enough. As shown in Table 2.4, this sample comprises 1644 observations, corresponding to about one fourth of the overall sample. Moreover, it encompasses a rather homogenous set of lenders, typically the largest players of the banking industry. As such, our analysis may underestimate the relevance of supply factors as a determinant of mortgage choice. For instance, it could be the case that large banks can more easily access financial markets to buy protection against interest rate risk or to raise long-term funds at fixed rate via covered bonds. If this is the case, focusing only on cross-border banks may lead to neglect part of the role played by supply conditions. To tackle this issue we conduct an exercise that requires a minor departure from our empirical setup. In particular, we consider time invariant country fixed effects and banking group fixed effects to capture demand and supply factors, respectively. In this way we are able to estimate similar regressions to those in Table 2.4, but run on the entire sample. We start with the specification shown in model 1 of Table 2.5 including only time dummies, which turn out to explain only a negligible portion of the total variation in the dependent variable (3%). Broadly speaking, this suggests that, in our sample, the cross section is a much more important dimension than the time series. Interestingly, by simply plugging country fixed effects, the R^2 raises to a surprising 70%. Model 3 displays instead the equation where the share of FRMs is regressed just on the set of banking group fixed effects. Despite the fact that these are largely collinear with the set of country fixed effects and significantly more granular,[19] the coefficient of determination, not only does not change, but actually

not driven by a higher number of dummy variables for one of the two groups. Additionally, 147 out of 360 month-country dummies are omitted because of collinearity, while no month-banking group dummy is omitted. Notwithstanding of that, month-country fixed effects have a higher explanatory power than month-banking group fixed effects.

[19]The two sets of fixed effects coincide in all observations related to banking groups operating only in one

slightly diminishes (69%) with respect to model 2. When we combine country dummies and bank dummies, as in model 4, we are able to explain almost 78% of the variation in the share. Using a Shorrocks-Shapely decomposition of the R^2, we find that country fixed effects exhibit a higher explanatory power than banking group fixed effects. The same applies in the two corresponding specifications also including month fixed effects, although, by construction, the R^2 raises somewhat. These considerations corroborate our conclusions drawn on the subsample of cross-border banks, emphasizing the role played by demand factors. As a further exercise, Table A2.1 in the Appendix shows the results of regressions including time invariant fixed effects run on the subsample of cross-border banking groups. Again, the role of time dummies is rather limited. Country fixed effects capture a sizable part of the variation in the share of FRMs, while banking group fixed effects have a much smaller explanatory power, as in Table 2.4.

[Insert Table 2.5 here]

2.5.3 Advanced Model

Regressions reported in previous tables provide a useful breakdown of the contribution of demand and supply factors in explaining the share of FRMs. This breakdown is powerful, as it relies on reasonable identifying assumptions. However, its main limitation is that it consists in a mere statistical decomposition, which prevents from providing a meaningful economic interpretation. In particular, as discussed earlier, our results suggest that demand factors play a prominent role, but these may include a rather heterogeneous set of borrower-specific characteristics. The normative conclusions may be quite different depending on what is the actual driver. We tackle this issue by adopting a hybrid approach. As in equation 2.1, we use month-banking group fixed effects to control for supply conditions. However, instead of introducing time varying country fixed effects to capture the demand, we directly model country-specific factors by including a set of variables suggested in the existing literature plus a novel variable. In particular, we consider the following variables: financial literacy, indebtedness, gross disposable income per capita, historical volatility of inflation, correlation between unemployment and the short-term interest rate, outstanding

country, which represent the vast majority of the sample. Moreover, the dataset includes 73 banking groups as opposed to only 12 countries.

amount of mortgage covered bonds to gross domestic product (GDP) and outstanding amount of residential mortgage-backed securities (RMBS) to GDP.

Our measure of *Financial Literacy* is obtained from the S&P Global FinLit Survey performed in 2014. The survey is based on interviews with more than 150000 adults in over 140 countries. It provides information on the degree of knowledge of four basic concepts in finance: risk diversification, inflation, numeracy and interest compounding. Financial literacy is measured as the percentage of 3 out of 4 answers correctly given by adults interviewed in each country. Table 2.7 and Figure A2.1 in the Appendix show that the level of financial education increases as we move from southern countries to northern countries. Financial literacy may have two opposite effects on the choice of FRMs versus ARMs. On the one hand, more educated borrowers understand that, unconditionally, a FRM is more expensive than an ARM and, hence, they are more likely to select an ARM (Agarwal et al., 2010; Gathergood and Weber, 2017). On the other hand, these educated borrowers may be more willing to choose a FRM, as they are aware of the risks related to the uncertain stream of payments of an ARM (Fornero et al., 2011).

To measure households' *Indebtedness* we use the ratio of total outstanding debt as percentage of gross disposable income provided by the OECD on a quarterly frequency. Table 2.7 and Figure A2.1 in the Appendix displays important differences in the level of households' indebtedness across countries. We consider the indebtedness ratio as a suitable proxy for households' income risk bearing capacity over the duration of the mortgage. Consistently with Campbell and Cocco (2003) and Fornero et al. (2011), we expect this ratio to have a positive effect on the share of FRMs.

As a measure of *Real Disposable Income Per Capita* we use the gross disposable income (adjusted for social transfers in kind) of households (and non-profit institutions serving households) expressed in purchasing power standard (PPS) per inhabitant, obtained from Eurostat on an annual basis. Table 2.7 and Figure A2.1 in the Appendix show a marked heterogeneity in households' real disposable income across countries over our sample period. The effect of disposable income on mortgage choice is rather ambiguous. It can capture either a current costs minimization effect (Campbell and Cocco, 2003), or an income risk bearing capacity effect (Ehrmann and Ziegelmeyer, 2017). If the first is prevalent, households with low income are more likely to select an ARM in order to minimize the current payment required by the loan. On the contrary, if the latter dominates, borrowers with low income are more prone to choose a FRM, because they may be concerned of not being able to face the future stream of payments required from an adjustable rate loan.

It is recognized in the literature that the unemployment rate plays a role in mortgage choice as well. For example, Ehrmann and Ziegelmeyer (2017) include among demand conditions the unemployment rate and its volatility, mainly as proxy for current and expected income. We believe that the unemployment rate is an important country demand factor, but we are aware that it may have opposite effects depending on whether households are mainly focused on current costs minimization or future income risk reduction.

A related aspect which has not been emphasized so far is that borrowers choosing between FRMs and ARMs should care, not only about the expected evolution in labor market conditions, but also about how unemployment will correlate with the level of interest rates. Risk-averse households expecting to be unemployed in a context of low interest rates tend to prefer, everything else equal, an ARM, as this implies a higher degree of consumption smoothing (mortgage installments decrease when income goes down and vice versa). Guren et al. (2018) provide a theoretical support for this argument. Usually a crisis unfolds because of a aggregate shock to the demand, leading to a drop in income and inflation. In such situation interest rates decrease, due to a possible decrease in expected inflation and especially to the monetary policy reaction of the central bank. Guren et al. (2018) show that, if the central bank reduces interest rates in response to a aggregate shock, households should select an ARM rather than a FRM. If, instead, interest rates increase during a downturn, for example because of a aggregate shock to the supply, households should prefer a FRM.

In light of that, the correlation between interest rates and unemployment depends on different factors including the slope of the Phillips curve and the monetary policy rule adopted. A full discussion of these aspects is clearly outside the scope of this paper. Here we limit ourselves to highlight that whenever such correlation is negative, the mortgage contract providing more protection against income fluctuations is, somewhat counterintuitively, the ARM and the insurance motive attached to it is stronger the smaller the correlation. We postulate that households make their expectations looking at the past. Then, to capture this effect we introduce a novel variable, namely the correlation between unemployment and the short-term interest rate.

We calculate *ρ(Unemployment, Short-term IR)* as the realized correlation between the unemployment rate and a short-term interest rate,[20] relying on a rolling window approach with a window of 7 years. We opt for a window of 7 years for two reasons: First, we assume that households make long-term expectations;[21] second, we make sure that, at the beginning of our sample period in 2007, we measure the correlation between these two variables after the introduction of the euro.[22] Table 2.7 and Figure A2.2 in the Appendix show that the correlation between unemployment and the short-term interest rate is negative in most countries over our sample period. This suggests that in periods of economic growth unemployment is low and the short-term interest rate is high as a result of a tight monetary policy aimed at containing inflation. Conversely, in bad times, as the recent double-dip European recession, unemployment is high and the short-term interest rate is low due to an expansionary monetary policy. Nevertheless, there are some exceptions. For example Germany exhibits a positive correlation from the end of 2010. The reason is that in 2009 the unemployment rate in Germany started to decrease, revealing a substantial improvement in economic fundamentals.[23]

We include as an indicator of the macroeconomic history of a country the volatility of the inflation rate over a period of 30 years prior to the introduction of the euro. We calculate *Historical Inflation Volatility* as the realized standard deviation of the monthly month-on-month inflation rate during the period 1970-1999 expressed in percentage points.[24] As in Campbell (2012), we estimate our measure on a pre-euro period in order to emphasize

[20]Data on short-term interest rates are retrieved from the OECD. For euro area countries the 3-month European Interbank Offer Rate is used from the date the country joined the euro. For the other countries the short-term interest rate is either the 3-month interbank offer rate or the yield on short-term Treasury bills, Certificates of Deposits or similar instruments with a maturity of three months.

[21]Usually long-term expectations have an horizon of at least five years (ECB, 2016, 2017).

[22]In this way we ensure that households expectations are made taking into account that monetary policy is defined by the ECB for the entire euro area. This clearly implies that we estimate the correlation between unemployment and short term interest rate having the same short-term interest rate for all countries (with the only exception of Greece, Latvia and Slovenia before their access to the euro area respectively in 2001, 2014 and 2007).

[23]This reflects, in turn, a flight-to-quality episode in the context of a monetary union. When economic conditions worsens due, as for example in the recent past, to a global financial crisis, policy rates go down to the same extent for every economy in the monetary union, but flight to quality makes unemployment raise more in peripheral countries. This also can explain why Germany is an outlier.

[24]Because of a lack in the available data, the historical volatility of inflation is computed over the period 1991-1999 for Latvia and 1980-1999 for Slovenia.

differences across countries. In Table 2.7 and Figure A2.1 in the Appendix we see that the periphery economies of the eurozone have experienced a substantial higher inflation volatility than central countries. High volatility of inflation is related to a higher share of ARMs. In order to understand why, the following considerations can be made. As mentioned above, when the correlation between unemployment and the short-term interest rate is negative, ARMs provide higher protection to borrowers. Economies where mortgages are predominantly at adjustable rate tend to be characterized by both a higher historical volatility of inflation and a larger, in magnitude, (negative) correlation between unemployment and the short-term interest rate, at least if compared to Germany (Table 2.7 and Figure A2.2 in the Appendix). Therefore, in these economies, the insurance provided by an ARM tends to be large and both factors, high inflation risk and a large, in magnitude, (negative) unemployment-interest rate correlation, contribute to it. Alternatively, Campbell (2012) and Badarinza et al. (2017) point out that a high volatility of inflation leads banks to set the interest rate on fixed rate loans at a relatively high level to protect them from inflation risk.[25] As a consequence, households are less willing to select a FRM. The fact that countries with a history of high inflation volatility still exhibit a prevalence of ARMs even after the introduction of the euro can only be interpreted as evidence of a sticky demand, suggesting that households tend to select the type of mortgage they are more familiar with (Campbell, 2012; Badarinza et al., 2017).

We label the variables listed so far as pure demand factors, as they relate to strictly specific households' characteristics. We take into account also two additional variables, namely *Outstanding Covered Bonds to GDP* and *Outstanding RMBS to GDP*. These are aimed to capture borrowers' characteristics that make mortgages extended locally suitable to back covered bonds or asset-backed securities. In principle these variables could capture both demand and supply factors. On the one hand, they can capture the reliance of banks on such funding instruments, highlighting an effect on the supply of FRMs. On the other hand, they can capture the eligibility of a mortgage issued locally to be used to secure covered bonds and mortgage-backed securities, assessing an effect on the demand of FRMs.[26] Nev-

[25]In particular, high inflation volatility entails a high cost of the prepayment option embedded in the interest rate charged on a FRM.

[26]For example, covered bonds regulations in most European countries specify that only mortgages having a loan-to-value below a certain threshold are eligible to be used as collateral for covered bonds (ECBC Covered Bond Comparative Database; ECB, 2008; ECBC, 2016).

ertheless, in our analysis these variables are mainly used to explain the demand, as supply conditions are captured by time-varying banking group fixed effects. We retrieve annual data on outstanding covered bonds from the European Covered Bond Council (ECBC). Our variable is the average over the last four years of the outstanding amount of mortgage covered bonds issued in a given country as percentage of GDP. Table 2.7 and Figure A2.2 in the Appendix show that mortgage covered bonds are particularly popular in Portugal and Spain. As for residential mortgage-backed securities, we get quarterly data from the Securities Industries and Financial Markets Association (SIFMA). Our variable is the average over the last four quarters of the outstanding amount of RMBS by country of collateral scaled by GDP. Table 2.7 and Figure A2.2 in the Appendix show that RMBS are common in the Netherlands and Portugal. Table 2.6 summarizes all the explanatory variables that we use to model country-specific factors, whilst Table 2.7 reports basic statistics.

[Insert Table 2.6 & Table 2.7 here]

Table 2.8 displays the estimates of the regressions including country specific explanatory variables and month-banking group fixed effects. In order to make sure that our regressors are predetermined, we include lagged values for those variables that are available on a lower frequency than monthly.[27] Given the different nature of the two groups of variables that we take into account, we first consider those capturing pure demand only and then integrate with the other country demand factors. Model 1 shows the results for the specification including pure demand factors only. We find a negative and significant coefficient for Real Disposable Income Per Capita in line with Ehrmann and Ziegelmeyer (2017). They interpret this finding with the view that households with higher income are more prone to select an adjustable rate loan, as they can comfortably face the income risk related to the uncertain stream of payments of an ARM. At the same time, and unlike what will be documented for Historical Inflation Volatility and ρ(Unemployment, Short-term IR), this finding is not robust to alternative specifications and should be considered with caution.

We find a negative and significant coefficient for the Historical Inflation Volatility, which confirms our priors. Our result is consistent with that of Campbell (2012) and Badarinza et al. (2017), showing that households' are more likely to select the type of loan they are more used to. An alternative explanation is that higher inflation risk entails a higher insurance

[27]These are all the explanatory variables except for ρ(Unemployment, Short-term IR).

motive attached to an ARM. As expected, the sign of the coefficients for Financial Literacy, Indebtedness and ρ(Unemployment, Short-term IR) are, respectively, negative, positive and positive, but neither of the three is statistically significant.

In model 2 we extend the previous specification by adding the two additional demand factors: Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. The sign and statistical significance of the pure demand regressors remains unchanged if compared to model 1, with the exception of ρ(Unemployment, Short-term IR).

The coefficient of ρ(Unemployment, Short-term IR) turns out to be positive and statistically significant, corroborating our view that the smaller such correlation, the stronger the insurance protection provided by an ARM. This suggests that households actually make expectations on what would be the macroeconomic environment in which a labor shock may occur. In particular, households that expect to be unemployed in a context of low interest rates are more willing to select an ARM, while households that envisage to be unemployed in a context of high interest rates, are more prone to choose a FRM. This result confirms the theoretical prediction of Guren et al. (2018).

The coefficients of the two additional variables are positive, but they result not to be statistically significant. To have a reliable basis for inference, both in model 1 and in model 2, we rely on standard errors clustered by country and quarter.[28] To tackle the issue that we may have few clusters, we adopt a small-sample correction for both standard errors and test statistics, as suggested by Cameron et al. (2011), and Cameron and Miller (2015). With such severe double clustering, Financial Literacy, Indebtedness and Outstanding Covered Bonds to GDP are not statistically significant.[29] However, most of them will recover significance in an alternative specification that overcomes the possible biases arising in this context, where we try to explain country demand factors relying on a sample with a heterogeneous coverage of banks across countries (analysis presented below).

In this type of exercise, we effectively control for bank supply conditions, but we cannot

[28]In Table A2.3 in the Appendix we show the evidence that lead us to adopt this type of clustering. By clustering at progressively higher levels in the two dimensions of our panel, we detect a substantial serial correlation and a less pronounced, but not negligible, cross correlation. This is why we decide to cluster at both the country and the quarter level. These two levels of clustering have been selected according to the procedure suggested by Petersen (2009), Cameron et al. (2011), and Cameron and Miller (2015).

[29]Table A2.2 in the Appendix shows that these variables are statistically significant when standard errors are not adjusted.

be entirely sure to capture at all country demand factors. We have relied on an exhaustive survey of existing papers in order to select a complete set of explanatory variables and we have actually enhanced it by introducing an additional (and novel) variable, i.e., ρ(Unemployment, Short-term IR). Nonetheless, we are aware that additional or alternative measures could be relevant in this setup. In order to assess whether our selection is reliable and comprehensive enough, we compare the quality of the fit obtained with the specification in model 2 with that obtained by replacing the explanatory variables with month-country fixed effects, but run on the sample used in model 2.[30] As shown in model 3, the latter amounts to 85% and represents the upper bound that can be reachable by including any arbitrarily large set of country-specific variables. The R^2 obtained by simply using our selection of seven regressors results to be quite close (79%).

Finally, Table A2.3 in the Appendix displays the results when adopting all possible alternative choices for double clustering of standard errors. Results are virtually unchanged.

[Insert Table 2.8 here]

2.5.4 Two-Stage Model

The results exposed so far provide useful insights on the determinants of the wide cross-country heterogeneity in the share of FRMs. Our findings suggest a prominent role for country demand factors, with a special emphasis on Real Disposable Income Per Capita, Historical Inflation Volatility and ρ(Unemployment, Short-term IR). Nevertheless, our sample is characterized by important differences in the number of banks operating in each country. As a consequence, we may wonder whether these results fully explain the mechanism behind the heterogeneity across countries, or rather they are driven by those countries that are more represented in our sample. In order to guarantee that we draw conclusions giving an equal weight to the observations pertaining to each country, we adopt a two stage approach, as in Ongena and Smith (2000). In particular, we regress the estimated coefficients of the month-country fixed effects in the full specification of equation 2.1 on our set of explanatory variables.[31] Unfortunately, 147 out of 393 month-country dummies in model 3 of

[30]Model 3 of Table 2.8 is equivalent to model 3 of Table 2.4, with the only difference that, in the former, the regression is run over a smaller sample to make it comparable to model 2 of Table 2.8. This is necessary because some of the regressors in model 2 of Table 2.8 are not available over some time periods.

[31]To perform the second stage regression we only need that the estimated coefficients of the month-county fixed effects are unbiased. We argue that this condition is satisfied as the time varying country fixed effects

Table 2.4 are omitted because of collinearity. As a consequence, performing the second stage regression with only 246 dependent variables would prevent us to get reliable results. To circumvent this issue, we estimate a similar regression to the one of equation 2.1, in which we substitute month-banking group fixed effects with quarter-banking group fixed effects. In this way, we are able to estimate 344 out of 393 month-country dummies and to perform the second stage regression accordingly. To be more specific, our two-stage regression looks as follows:

$$share(b,c,t) = \alpha(c,t) + \beta(h(b),t) + \varepsilon(b,c,t) \tag{2.2}$$

$$\hat{\alpha}(c,t) = x'(c,t)\gamma + \upsilon(c,t) \tag{2.3}$$

The terms $\beta(h(b),t)$ represent quarter-banking group fixed effects, while $x'(c,t)$ denotes the vector of explanatory variables capturing demand conditions.

Table 2.9 reports the results of the first stage and the second stage regressions. Model 1 consists in the regression in which we include month-country fixed effects and quarter-banking group fixed effects. To check if by substituting month-country fixed effects with our seven regressors we alter the findings exposed in Table 2.8, we include model 2 and everything remains virtually unchanged. In models 3-4, the coefficients of month-country fixed effects estimated by running model 1 are regressed over the set of explanatory variables capturing demand conditions. Model 3 includes pure demand factors only. As before, we find a negative and significant coefficient for Historical Inflation Volatility. Model 4 extends the preceding including all set of regressors. Historical Inflation Volatility maintains its sign and significance. Similarly to Table 2.8, ρ(Unemployment, Short-term IR) exhibits a positive and significant coefficient. As for the other variables, we detect important differences with respect to model 2 of Table 2.8. The coefficients of Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP are both positive and significant, suggesting that these country demand factors actually matter. In countries where the characteristics of borrowers ease the issuance of covered bonds and asset-backed securities FRMs are relatively more appealing. Statistically significant is also the coefficient of Financial Literacy. The negative sign implies that financially educated households are more willing to select an ARM, as they are able to understand that, unconditionally, an ARM is cheaper than a

and banking group fixed effects included in the first stage regression span all the possible factors determining the dependent variable.

FRM. In contrast to previous results, Real Disposable Income Per Capita loses its significance. Tables A2.4-A2.5 in the Appendix report the results of models 2-4 when adopting all possible alternative choices for double clustering of standard errors. Results are virtually unchanged. Table A2.6 displays similar findings when the share of FRMs is decomposed in month-country fixed effects and year-banking group fixed effects.

[Insert Table 2.9 here]

To obtain relevant normative insights, we do not limit ourselves to merely identifying the country demand factors that play a role in mortgage choice, but we also provide an economic assessment of their magnitude. Table 2.10 reports the magnitude effects of the seven variables included in model 4 of Table 2.9. Focusing the attention on those that are statistically significant, we find that the Historical Inflation Volatility exhibits the strongest effect. One standard deviation increase leads to a decrease of 59 percentage points in the average share of FRMs per country cleaned of variation due to bank supply factors. Sizable is also the effect of Financial Literacy. A rise of one standard deviation corresponds to a drop of 42 percentage points in the average share of FRMs per country ascribable to demand factors. Moreover, a one standard deviation increase in Outstanding Covered Bonds to GDP and in Outstanding RMBS to GDP determines a rise, respectively, of 32 and 17 percentage points in the dependent variable. Finally, a one standard deviation increase in ρ(Unemployment, Short-term IR) leads to a rise of 14 percentage points in the average share of FRMs per country left unexplained by bank supply factors.

[Insert Table 2.10 here]

2.5.5 Time Variation

Some useful indications can be obtained by exploring more closely the variation across time of the share of FRMs. As noted in Figure 2.1, for those countries in which the share of FRMs changes over time, the variability seems to be related to the spread between FRMs and ARMs interest rates. Since the term spread is a component of the spread between the interest rate applied on fixed rate and adjustable rate loans, the time variation in the share is related to the term spread as well. We aim to investigate whether the sensitivity of the share of FRMs to the term spread is mainly driven by the demand or the supply. To this end we perform the following type of regression:

$$\begin{aligned} share(b,c,t) =& \alpha(c) + \alpha(c) \times tspread(t) \\ & + \beta(h(b)) + \beta(h(b)) \times tspread(t) + \varepsilon(b,c,t) \end{aligned} \tag{2.4}$$

The terms $\alpha(c)$ represent country fixed effects, $\beta(h(b))$ denotes banking group fixed effects and $tspread(t)$ is the term spread at time t.

In this model, country fixed effects and banking group fixed effects capture the average level of the share for each country and each banking group. Their interactions with the term spread capture, instead, the sensitivity (slope) of each country and each banking group to changes in the term spread. This regression allows us to model the time variation in the share of FRMs using the term spread and assuming that the relation between these two is linear. As before, to disentangle shifts in demand from shifts in supply, we focus the attention on cross-border banking groups.

It is important to stress that, differently from other studies, we regress the share of FRMs on the term spread rather than the spread between FRMs and ARMs interest rates, as we want to draw causal inference. While the former can be considered to a large extent exogenous, the latter is inherently endogenous. Indeed, the spread between FRMs and ARMs interest rates is simultaneously determined with the quantities of FRMs and ARMs extended in equilibrium.

In estimating this model we use the term spread computed at the European level as the difference between the 10-year Interest Rate Swap rate and the 3-month Overnight Index Swap rate. We adopt this measure for the slope of the yield curve rather than the term spread for each country obtained on the basis of the respective government bonds. The reason is that, especially for those country that were more affected by the sovereign debt crisis, the interest rate charged on FRMs is more closely related to the 10-year Interest Rate Swap rate rather than the yield on 10-year government bonds. This can be explained by the fact that, during most of the time period under analysis, sovereign default risk in several countries was sensibly higher than credit risk associated with local mortgages.

Table 2.11 reports six different specifications. Model 1 includes country fixed effects only, while model 4 includes both country fixed effects and their interaction with the term spread. Country fixed effects explain alone 58% of the variation in the share of FRMs. When we add the interaction of country fixed effects with the term spread the coefficient of determination rises to 66%. This value is quite far from the 84% achieved in our baseline model with month-country fixed effects. However, while in the baseline model we allow country fixed effects to vary in a discretionary way over time, in model 4 we constrain the share of fixed rate mortgages to evolve linearly with the term spread. Of course, since the share is bounded between 0 and 100, it is likely that this relation is not linear. In fact, if we add an additional interaction term with the term spread squared, we experience an increase in the R^2 (71%). So, we conclude that the term spread is able to explain the time variation in the share of FRMs and that the relation between these two is not perfectly linear. A similar argument applies also to the two specifications with banking group fixed effects, namely model 2 and model 5.

Consistently with the evidence in Figure 2.1, we find that most of the coefficients of the interaction terms in model 4 are negative and significant. However, the sensitivity of the share of FRMs to the term spread differs significantly across countries. In particular, Belgium, Greece, Italy, Luxembourg and Slovenia are those countries where the share of FRMs is more reactive to changes in the term spread.

We have already pointed out that changes in the term spread can shift both the demand and the supply. On the one hand, an increase in the term spread, driven by an increase in inflation risk, may lead banks to decrease the supply of fixed rate loans, by making them relatively more expensive than adjustable rate ones. On the other hand, a rise in the spread between FRMs and ARMs interest rates due to an increase in the term spread may induce households to reduce their demand for fixed rate loans, which could signal either some form of myopic behavior (households choose ARMs when the term spread is high because they tend to give too much importance to the first installments), as well as the presence of financial constraints (matched with expectations of an increase in income). To assess whether the demand or the supply is more sensitive to changes in the slope of the yield curve, we include a specification in which we interact both country fixed effects and banking group fixed effects with the term spread. Relying on the Shorrocks-Shapely decomposition, we are able to detect the contribution of each interaction to the R^2. Model 6 shows that the fraction of R^2 ascribable to the interaction between country fixed effects and the term spread is much higher than the fraction attributable to the other interaction. Thus, we conclude that changes in the slope of the yield curve shift mainly the demand.

[Insert Table 2.11 here]

2.6 Tobit Robustness Checks

The results exposed so far are obtained using linear regressions. Our dependent variable, the share of FRMs, is a percentage bounded between 0 and 100. Using a linear model in this setting leads to inconsistent estimates. For this reason, it should be more appropriate to use a censored Tobit model of the form:

$$y^* = x\beta + \varepsilon \tag{2.5}$$

$$y = \begin{cases} 0 & \text{if} \quad y^* < 0 \\ y^* & \text{if} \quad 0 \leq y^* \leq 100 \\ 100 & \text{if} \quad y^* > 100 \end{cases}$$

Nonetheless, most of our findings are drawn by comparing the coefficients of determination of different specifications. Unfortunately, Tobit models do not provide such measure. Alternative metrics known as pseudo-R^2 cannot be considered as meaningful as the coefficient of determination of linear models. Moreover, in the specifications where we model the demand relying on a set of explanatory variables, we control for bank supply conditions including month-banking group fixed effects. It is well known that nonlinear models with fixed effects suffer from the so called "incidental parameters problem" (Neyman and Scott, 1948; Lancaster, 2000). This implies that the maximum likelihood estimator (MLE) is inconsistent. Greene (2004a,b) shows that, for the specific case of Tobit models with fixed effects, the slope coefficients are slightly affected by the incidental parameters problem. However, the bias can be sizable for the disturbance variance, with clear implications also on the estimation of the marginal effects. Therefore, either using linear or nonlinear models, we have to deal with relevant issues that can produce unreliable results. In light of the fact that our sample includes only four observations where the share of FRMs is exactly equal to one of the two bounds,[32] we believe that the issue related to linear regression models is less severe and, hence, we rely on them to derive our main results. Nonetheless, we perform a set of Tobit robustness checks in order to test whether our findings are robust to nonlinear specifications.

[32] In these four observations the value of the share is equal to the upper bound 100.

We start by replicating Table 2.4 using a censored regression model with lower bound 0 and upper bound 100. We calculate the pseudo R^2 according to the methodology suggested in Wooldridge (2010). In particular, we computed it as the square of the correlation coefficient between the dependent variable and the estimate of $\mathbb{E}\left[y\,|x\right]$. Table A2.7 in the Appendix shows that, as before, month-country fixed effects explain a larger fraction of the variation in the dependent variable than month-banking group fixed effects. We extend our analysis also including Tobit models with lower bound 1 and upper bound 99, in order to check whether our findings are affected by a more restrictive censoring. Results are virtually unchanged.

Tables A2.8-A2.9 in the Appendix replicate Table 2.5. In both tables the pattern of the R^2 across the different specifications is equal to the one displayed in Table 2.5. This confirms the prominent role of country demand factors, even when considering the whole sample of banks. However, in this setting we are not able to perform a decomposition of the R^2 to get additional insights.

Table A2.10 in the Appendix shows the estimates of the censored regression models including country specific explanatory variables and month-banking group fixed effects. For each regressor we report the marginal effect of the censored variable $\mathbb{E}\left[y\,|x\right]$ at the sample means. Differently from Table 2.8, we cluster standard errors only by country, as the statistical software that we use does not allow to implement two-way clustering in the Tobit model that we employ. We consider this a minor limitation, as we detected a higher serial correlation than cross correlation in our data set. In the specifications with the full set of country variables, we find, as before, a negative and statistically significant coefficient for Real Disposable Income Per Capita and Historical Inflation Volatility, as well as a positive and statistically significant coefficient for ρ(Unemployment, Short-term IR).

As in the previous section, we improve our analysis making sure that we equally weight each country when explaining the cross-country heterogeneity in the share of FRMs. To this aim, we rely on a two-stage approach. In the first stage we perform a censored regression including month-country fixed effects and quarter-banking group fixed effects. In the second stage we regress the estimated coefficients of the month-country fixed effects, which correspond to the marginal effects of the latent variable y^*, on our set of explanatory variables. While in the first stage we use a Tobit model, in the second stage we employ a linear regression, as the dependent variable is not constrained between 0 and 100. Model

4 of Tables A2.11-A2.12 in the Appendix shows, as in Table 2.9, a negative and significant coefficient for Financial Literacy and Historical Inflation Volatility, as well as a positive and significant coefficient for ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP.

Finally, in Table A2.13 in the Appendix we investigate the time variation in the share of FRMs using censored regression models. As in Table 2.11, we find that, the sensitivity of the share of FRMs to the term spread is quite heterogeneous across countries. Moreover, the term spread captures an important fraction of the time variation in the dependent variable.

The Tobit robustness checks exposed above highlight that the results obtained using linear regression models are indeed robust to nonlinear specifications.

2.7 Empirical Analysis on the Spread

The quantity of FRMs and ARMs, as well as their interest rates, are simultaneously determined on the market by the interaction between demand and supply. No bank should be able to individually set the share of FRMs granted neither the price of FRMs and ARMs. If this is the case, the variation in the spread between FRMs and ARMs interest rates should be explained by the same factors driving the share of FRMs. We want to explore this possibility by performing the same set of reduced-form regressions exposed in Section 2.5 using this time as dependent variable the spread between FRMs and ARMs interest rates, henceforth abridged simply with "spread".

Models 1-3 of Table A2.14 in the Appendix displays three specifications in which the spread is regressed on, respectively, month-country fixed effects, month-banking group fixed effects and both sets of fixed effects jointly. Month-country fixed effects alone explain 60% of the variation in the spread, suggesting that, also in this case, country demand factors play a major role. Month-banking group fixed effects explain only 38% of the variation in the dependent variable, but the difference between the R^2 of model 1 and model 2 is smaller compared to what seen for the share of FRMs in Table 2.4. When taken together the two sets of fixed effects can explain 73% of the total variation in the spread. We conclude that also the spread is mainly driven by the demand, although here our model is somewhat less capable of describing the data. The supply plays a role as well and it seems to be slightly more relevant in explaining the spread than the share of FRMs.

The following step is to model month-country fixed effects with the selection of regressors that we used in Section 2.5. We expect that these explanatory variables have an effect also on the spread, but the relation should be of opposite sign with respect to the one observed in the analysis on the share of FRMs. To avoid possible distortions related to heterogeneous coverage of the dataset across countries (in terms of number of intermediaries) we directly look at the two-stage approach, as described above for quantities. Model 1 of Table A2.15 in the Appendix consists in the regression with month-country fixed effects and year-banking group fixed effects. We report the results of this specification including year-banking group fixed effects, instead of quarter-banking group fixed effects, because the results are not exactly the same under the two models. In light of that, we consider the specification with year-banking group fixed effects more reliable, as it allows us to perform the second stage regression having 381 out of 393 estimated coefficients of month-country fixed effects. As shown in model 4 of Table A2.15 in the Appendix, two factors turn out to be significant, at least when a two-way cluster by country and quarter is adopted, both with the expected sign: the ρ(Unemployment, Short-term IR) and the Outstanding RMBS to GDP. In general, the coefficients of all the explanatory variables are very little and sensibly lower than those displayed in Table 2.9. The weak effects of our regressors are hardly surprising though. In fact, as highlighted before, the cross-country variation in the spread is much lower than the variation in the share of FRMs across countries.

We extend our analysis looking at the time variation in the spread. Model 6 of Table A2.16 includes country fixed effects, banking group fixed effects, as well as their interaction with the term spread. The R^2 of this specification (58%) is relatively high but fifteen percentage points lower than the coefficient of determination of our baseline model with month-country fixed effects and month-banking group fixed effects (73%). As before, this suggests that the term spread is able to capture the time variation in the spread, but the relation with the dependent variable might be nonlinear. In Figure 2.1 we observed that the evolution of the spread over time is directly related to the evolution of the term spread. The positive and significant coefficients of the interactions between country fixed effects and the term spread confirm this evidence. As for the share of FRMs, in Belgium, Greece, Italy, Luxembourg and Slovenia the spread is more sensitive to changes in the term spread. The Shorrocks-Shapley decomposition of the R^2 of model 6 eventually corroborates that the term spread is mainly able to shift the demand, although the effect it exerts on the supply is slightly higher than what is detected in Table 2.11.

2.8 Conclusions

Using granular bank level information from 103 banks belonging to 73 different banking groups across twelve countries in the euro area, we provide a comprehensive analysis of the determinants of mortgage choice in the euro area. In particular, we investigate to what degree the wide cross-country heterogeneity in the share of fixed rate to total new mortgages is driven by differences in demand or supply conditions.

Relying on a prudent identification strategy, we are able to explore the role of country demand and bank supply factors in determining households' mortgage choice. Specifically, we assume that lending policies are set at the consolidated level and can disentangle demand from supply by comparing the lending patterns observed for the same cross-border banking group in different euro area economies, as well as the lending patterns observed across different cross-border banking groups operating in the same economy. Country demand conditions results to have a prominent role in driving the prevalence of mortgages extended at a fixed rate. In particular, they are able to explain almost 72% of the total variation of the share of fixed rate to total new mortgages observed in the sample.

Factors such as the historical volatility of inflation rates, the correlation between unemployment and the short-term interest rate, households' financial literacy, and the volume of outstanding mortgage covered bonds and mortgage-backed securities exhibit a high correlation with the estimated demand component of fixed rate mortgages, relative to adjustable rate ones.

A predominant role for demand factors is documented also when focusing on the sensitivity of the share of fixed rate mortgages to the slope of the yield curve, as well as when analyzing lending conditions, that is the spread between the interest rate on fixed rate mortgages and that on adjustable rate mortgages.

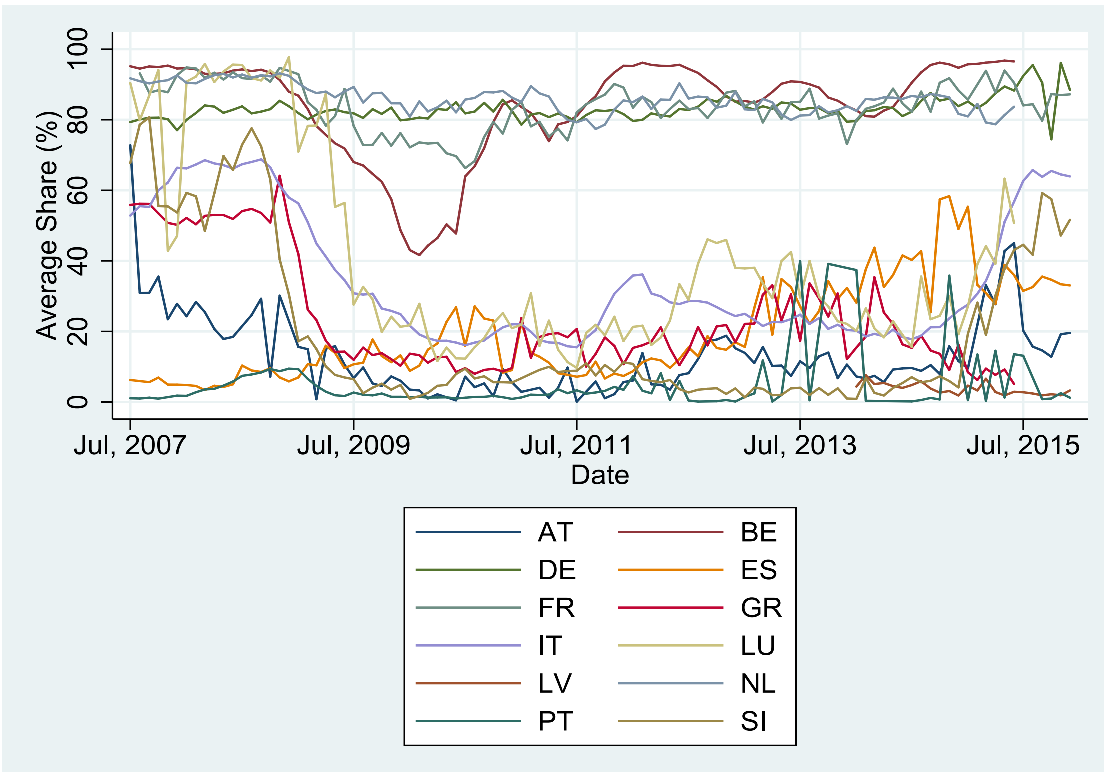
By showing the relevance of country demand factors, a policy implication of our analysis is that it would not make sense to try to influence the share of fixed rate mortgages by pressing banks to take on more duration risk. This would be ineffective and, presumably, even not desirable. Indeed, the heterogeneity in the share of fixed rate mortgages across economies seems to reflect an optimal allocation of interest rate risk, given the asynchronous business cycles and the expectations that monetary policy will operate in a way that stabilizes disposable income net of housing costs.

2.9 Figures

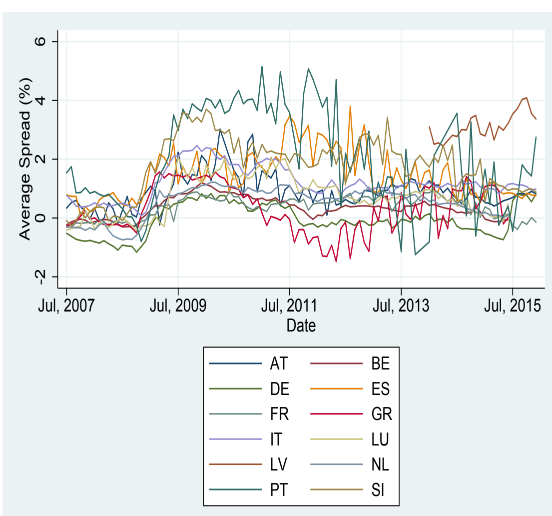
Figure 2.1: Share of FRMs and spread between FRMs ands ARM interest rates

Average share of FRMs (a), average spread between FRMs and ARMs interest rates (b), term spread computed as the difference between the 10-year Interest Rate Swap rate and the 3-month Overnight Index Swap rate (c).

a. Average share of FRMs



b. Average spread of FRMs-ARMs interest rates



c. Term spread

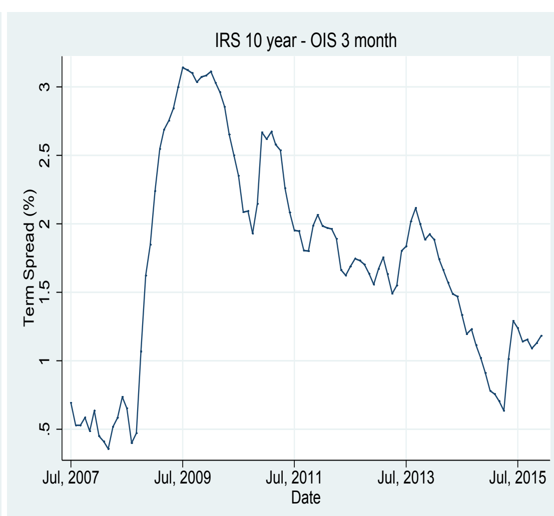
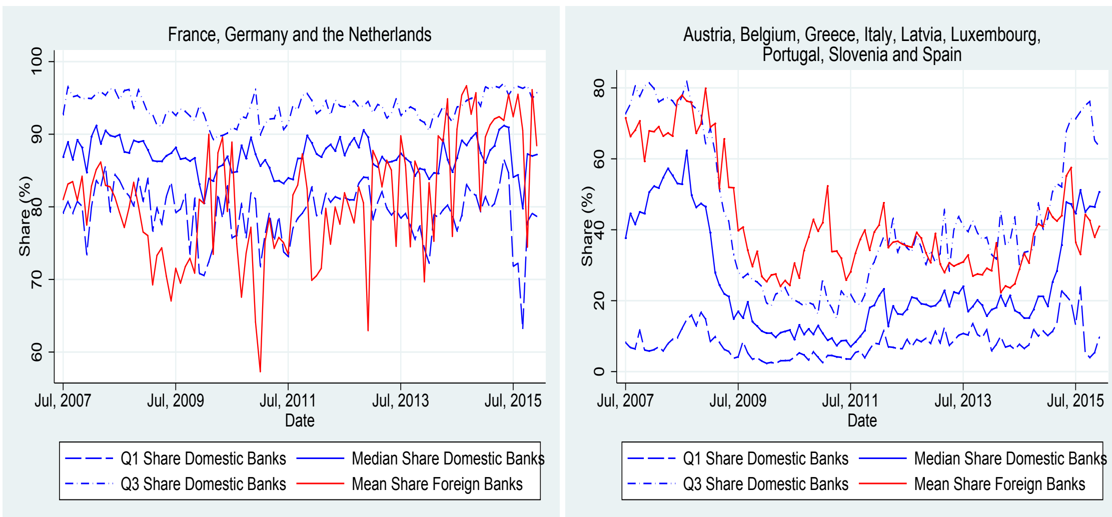


Figure 2.2: Share of FRMs for groups of countries

Share of FRMs of domestic banks and foreign banks for the first group (a) and the second group (b) of countries. Domestic banks are banks with a domestic bank holding. Foreign banks are banks with a foreign bank holding. The first group includes France, Germany and the Netherlands. The second group includes Austria, Belgium, Greece, Italy, Latvia, Luxembourg, Portugal, Slovenia and Spain. Q1 and Q3 stand for first quartile and third quartile, respectively.

a. Share of FRMs 1st group

b. Share of FRMs 2nd group



2.10 Tables

Table 2.1: Determinants of the Share of FRMs and/or the Probability of a FRM Choice Identified in the Literature

↑ and ↓ denote a positive and a negative effect, respectively. ***,** and * stand for statistical significance at 1%, 5% and 10%. 0 denotes a variable included in the specification which is not significant. Papers are indicated by numbers from 1 to 14 as follows: [1] Paiella and Pozzolo (2007), [2] Koijen et al. (2009), [3] Agarwal et al. (2010), [4] Fornero et al. (2011), [5] Campbell (2012), [6] Fuster and Vickery (2014), [7] Foà et al. (2015), [8] Badarinza et al. (2017), [9] Basten et al. (2017), [10] Ehrmann and Ziegelmeyer (2017), [11] Gathergood and Weber (2017), [12] Campbell and Cocco (2003), [13] Kirti (2017), [14] Guren et al. (2018). Statistical significance is not reported for the variables investigated by Agarwal et al. (2010) and Campbell (2012), as well as for the historical inflation volatility analysed by **?**, as they do not perform a regression analysis. The variables reported for Badarinza et al. (2017) refer to the analysis on 3-year (left) and 1-year (right) expectations of the future ARM rate.

		Empirical papers																
	Paper	[1]	[2]		[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]				This paper	
	Country	IT	US		US	IT	EU, US, CA	US	IT	EU, US, AU	CH	Euro Area	UK				Euro Area	
	Sample Years	1995-2004	1985-2006		2005-2007	2005-2008		1996-2009	2004-2010	1990-2013	2010-2013	<1980-2010	2013	Theoretical papers			2007-2015	
	Variable of interest	Prob. FRM	Prob. FRM	Share FRM	Prob. FRM	Prob. FRM	Share FRM	Prob. FRM	Prob. FRM	Share FRM	Prob. FRM	Prob. FRM	Prob. FRM	[12]	[13]	[14]	Our variable	Share FRM
	Age	↑**				0					↓***							
	Gender	0				0												
	Married	0																
	Children	↑**																
	Income	0										↓***		↑			Real disposable income per capita	0
	N. income recipients	0				0												
Borrower	Nondurable expenditures	↓**																
Variables	Financial wealth	0				0					↑***			↑				
Demand	Education	0				0												
	Financial literacy				↓	↑**							↓***				Financial literacy	↓**
	Type of employment	0				0												
	Mobility	0												↓				
	Risk aversion					0							0	↑				
	Volaility of labor income													↑				
	Other debt										↓**							
	Credit score		↑***															
	Duration											↓***						
	Mortgage to income					↑***							0	↑			Indebtedbeness	0
Loan	House price	↓***																
Variables	House price to income	↑***																
Demand/Supply	Debt service to income										0	↓***						
	Loan to value		↓***								↓**		↓**					
	Loan balance		↑***															
	Deposits to total liabilities								↑***									
	Deposits to total assets										↓**							
	Equity to total assets										↓*							
Bank	Floating rate liabilities														↓			
Variables	Bond spread								↓*									
Supply	Exposure to interest rate risk										↓***							
	Derivatives usage										0							
	Size										↑***							
	Competition	↓**																
	Securitisation							↑***	↑***								Outstanding RMBS to GDP	↑***
	Covered bonds																Outstanding covered bonds to GDP	↑***
	Spread FRM-ARM interest rates	↓***		↓***		↓***				↓***/0								
	FRM rate minus expected ARM rate		↓***	↓***					↓***	0/↓***								
	ARM interest rate	↓**																
Macroeconomic	Long-term interest rate	0	↑***	↓***								0		↓				
and	Term spread		↓***	0							↓***	↓***		↓			Term spread	↓*,**,***
Institutional	Inflation rate											↑**						
Variables	Inflation rate volatility											0						
Demand/Supply	Historical inflation volatility						↓			↓							Historical inflation volatility	↓***
	Unemployment rate											0						
	Unemployment rate volatility											↑**						
	GDP growth					↑*						↓***						
	GDP growth volatility											0						
	Correlation aggregate shocks short-term interest rate															↑	Correlation unemployment short-term interest rate	↑**

Table 2.2: Overview of Banks and Banking Groups, by Country

Country	Banks with a domestic bank holding	Banks with a foreign bank holding	Banks belonging to a cross-border banking group	Domestic banking groups	Cross-border banking groups
Germany	35	1	5	26	1
Italy	16	2	3	12	1
France	13	0	4	2	3
Spain	10	1	1	9	0
Austria	3	1	1	3	0
Slovenia	2	2	2	2	0
Belgium	3	1	1	3	0
Greece	4	0	0	4	0
The Netherlands	0	3	0	3	0
Portugal	3	0	0	3	0
Luxembourg	0	2	2	0	0
Latvia	1	0	0	1	0
Total	93	10	19	68	5

Table 2.3: Overview of the Share of FRMs and the Spread between FRMs and ARMs interest rates, by Country

		Share of FRMs (%)				**Spread FRMs - ARMs interest rates (%)**			
Country	N	Average	Median	Minimun	Maximum	Average	Median	Minimum	Maximum
Austria	223	12.66	7.74	0.09	74.30	0.90	0.84	-0.54	3.49
Belgium	377	84.17	91.48	21.10	99.99	0.34	0.35	-1.04	1.70
France	812	84.37	93.14	6.25	100.00	0.44	0.38	-4.75	3.47
Germany	2565	82.78	85.93	14.95	99.96	-0.09	-0.04	-3.34	2.67
Greece	261	26.73	17.05	0.26	88.71	0.34	0.50	-2.08	3.39
Italy	1614	33.77	25.90	0.17	98.15	1.21	1.15	-1.12	3.43
Latvia	24	3.74	3.28	1.85	7.57	3.12	3.05	2.45	4.09
Luxembourg	161	36.39	28.50	1.58	97.95	0.86	0.77	-0.42	2.49
Portugal	183	4.50	1.99	0.05	39.93	1.94	1.83	-1.75	6.08
Slovenia	254	16.78	4.86	0.09	94.10	1.85	1.91	-0.33	4.13
Spain	605	19.10	8.37	0.09	90.84	1.49	0.93	-1.57	7.47
The Netherlands	248	86.30	86.01	71.43	98.47	0.50	0.72	-1.14	1.70
Sample	7327	57.44	71.26	0.05	100.00	0.62	0.55	-4.75	7.47

Table 2.4: Baseline model

Fixed effects decomposition of the share of FRMs. Sample: cross-border banking groups. Model: linear. Dependent variable: share of FRMs. Standard errors: not adjusted. Shorrocks-Shapely decomposition of the R^2 in model 3.

	(1)	(2)	(3)	(4)
Month-country FE	YES	-	YES	YES
Month-banking group FE	-	YES	YES	YES
Bank FE	-	-	-	YES
N	1644	1644	1644	1644
R^2	0.843	0.319	0.908	0.973
Adjusted R^2	0.731	0.038	0.746	0.924
R^2 month-country FE			0.716	
R^2 month-banking group FE			0.191	
F-test statistic	7.493***	1.137**	5.616***	19.897***
degrees of freedom	(688,956)	(480,1164)	(1046,598)	(1057,587)

* p<0.10, ** p<0.05, *** p<0.01

Table 2.5: Fixed effects decomposition with time invariant fixed effects

Fixed effects decomposition of the share of FRMs. Sample: all banks. Model: linear. Dependent variable: share of FRMs. Standard errors: not adjusted. Shorrocks-Shapely decomposition of the R^2 in model 4 and 5.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Month FE	YES	-	-	-	YES	YES	YES
Country FE	-	YES	-	YES	YES	-	YES
Banking group FE	-	-	YES	YES	-	YES	YES
N	7327	7327	7327	7327	7327	7327	7327
R^2	0.026	0.697	0.687	0.779	0.735	0.724	0.818
Adjusted R^2	0.012	0.696	0.684	0.776	0.730	0.717	0.813
R^2 month FE							0.034
R^2 country FE				0.394			0.397
R^2 banking group FE				0.385			0.387
F-test statistic	1.879***	1528.181***	221.414***	323.015***	178.15***	108.512***	178.425***
degrees of freedom	(102,7225)	(12,7315)	(73,7254)	(80,7247)	(113,7214)	(174,7153)	(181,7146)

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 2.6: Description of Country Demand Variables

Variable	Description
Financial Literacy	Percentage of 3 out of 4 answers correct given by adults interviewed in each country, as results from the S&P Global FinLit Survey.
Indebtedness	Ratio of total outstanding debt as percentage of gross disposable income provided by the OECD on a quarterly frequency. Data are missing for Latvia and Luxembourg, and partially available for Greece, Italy and the Netherlands.
Real Disposable Income Per Capita	Gross disposable income (adjusted for social transfers in kind) of households (and non-profit institutions serving households) expressed in purchasing power standard (PPS) per inhabitant, obtained from Eurostat on an annual basis. Data are missing for Luxembourg.
Historical Inflation Volatility	Realized standard deviation of the monthly month-on-month inflation rate during the period 1970-1999. Because of a lack in the available data, Historical Inflation Volatility is computed over the period 1991-1999 for Latvia and 1980-1999 for Slovenia. Monthly data on the inflation rate are retrieved from the OECD.
ρ(Unemployment, Short-term IR)	Realized correlation between the unemployment rate and the short-term interest rate, calculated on a rolling window approach with a window of 7 years. Monthly data on unemployment rates and short-term interest rates are retrieved from the OECD.
Outstanding Covered Bonds to GDP	Average over the last four years of the amount outstanding of mortgage covered bonds as percentage of GDP. Annual data on outstanding covered bonds are retrieved from the European Covered Bond Council (ECBC). Data are missing for Slovenia.
Outstanding RMBS to GDP	Average over the last four quarters of the amount outstanding of RMBS as percentage of GDP. Quarterly data on outstanding residential mortgage-backed securities are retrieved from the Securities Industries and Financial Markets Association (SIFMA). Data are missing for Latvia, Luxembourg and Slovenia and not available for all other countries in 2007.

Table 2.7: Overview of Country Demand Variables, by Country

	Financial literacy (%)		**Indebtedness (%)**		**Real disposable income per capita (PPS)**		**Historical inflation volatility (%)**		**Correlation unemployment short-term interest rate**		**Outstanding covered bonds bonds to GDP (%)**		**Outstanding RMBS to GDP (%)**	
	Avg.	Std. Dev.	Avg.	Std. Dev.	Avg.	Std. Dev.	Avg.	Std. Dev.	Avg.	Std. Dev.	Avg.	Std. Dev.	Avg.	Std. Dev.
Austria	53.00	-	85.73	1.70	24746.56	1065.49	44.68	-	-0.63	0.13	3.40	1.96	0.66	0.10
Belgium	55.00	-	91.73	7.53	22621.89	1140.56	38.38	-	-0.64	0.15	0.50	0.81	14.00	4.58
France	52.00	-	98.36	3.77	23103.56	1072.30	39.73	-	-0.89	0.09	6.71	2.77	0.88	0.49
Germany	66.00	-	89.45	3.05	25258.56	1862.56	32.39	-	0.04	0.59	8.39	0.83	0.63	0.17
Greece	45.00	-	104.29	3.26	16999.33	1870.32	152.30	-	-0.69	0.19	4.91	3.70	3.08	0.56
Italy	37.00	-	83.37	0.78	20926.00	355.69	57.32	-	-0.70	0.24	2.74	2.88	6.17	1.77
Latvia	48.00	-	-	-	12367.50	549.42	1015.10	-	0.10	0.22	0.02	0.03	-	-
Luxembourg	53.00	-	-	-	-	-	40.46	-	-0.36	0.16	0.13	0.14	-	-
Portugal	26.00	-	134.46	4.59	16340.56	415.05	146.49	-	-0.47	0.36	12.28	7.22	17.56	3.14
Slovenia	44.00	-	52.70	2.22	16011.22	466.54	859.02	-	-0.40	0.50	-	-	-	-
Spain	49.00	-	132.03	8.35	18220.56	404.43	70.06	-	-0.66	0.26	28.92	5.56	13.64	2.41
The Netherlands	66.00	-	263.55	3.53	22865.44	321.71	44.36	-	-0.70	0.09	5.41	3.10	35.85	8.30
Sample	49.50	11.15	109.97	50.93	20528.02	3686.80	211.69	342.87	-0.54	0.39	7.18	8.68	10.09	11.35

Table 2.8: Advanced model

Sample: cross-border banking groups. Model: linear. Dependent variable: share of FRMs. Explanatory variables: Financial Literacy, Indebtedness, Real Disposable Income Per Capita, Historical Inflation Volatility, ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. Standard errors: two-way clustered by country and quarter for model 1-2, not adjusted for model 3.

	(1)	(2)	(3)
Financial Literacy	-0.509	-1.689	
	(1.84)	(1.47)	
Indebtedness	0.835	0.602	
	(0.48)	(0.55)	
Real Disposable Income Per Capita	-0.014***	-0.012**	
	(0.00)	(0.00)	
Historical Inflation Volatility	-5.221***	-5.799***	
	(1.14)	(0.68)	
ρ (**Unemployment, Short-term IR**)	20.473	24.170**	
	(11.23)	(8.18)	
Outstanding Covered Bonds to GDP		1.430	
		(1.31)	
Outstanding RMBS to GDP		0.319	
		(0.88)	
Month-banking group FE	YES	YES	YES
Month-country FE	-	-	YES
Two-way cluster	*country, quarter*	*country, quarter*	-
N	1085	1085	1085
R^2	0.785	0.789	0.852
Adjusted R^2	0.677	0.682	0.666
F-test statistic regressors	276.015***	-	
degrees of freedom	(5,5)	-	
F-test statistic regressors pure demand		158.955***	
degrees of freedom		(5,5)	
F-test statistic regressors institutional demand		1.829	
degrees of freedom		(2,5)	
F-test statistic fixed effects			4.572***
degrees of freedom			(606,479)

* p<0.10, ** p<0.05, *** p<0.01

Table 2.9: Two-stage model

First stage regressions including: month-country fixed effects and quarter-banking group fixed effects (1), all explanatory variables and quarter-banking group fixed effects (2). In the first stage regressions the dependent variable is the share of FRMs. Second stage regressions of the estimated coefficients of month-country fixed effects in (1) on: pure demand explanatory variables (3) and all explanatory variables (4). Sample: cross-border banking groups. Model: linear. Explanatory variables: Financial Literacy, Indebtedness, Real Disposable Income Per Capita, Historical Inflation Volatility, ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. Standard errors: not adjusted for model 1, two-way clustered by country and quarter for model 2-4.

	1ST STAGE		***2ND STAGE***	
	(1)	(2)	(3)	(4)
Financial Literacy		-1.634	-2.693	-5.386**
		(1.37)	(2.26)	(1.72)
Indebtedness		0.586	1.558	0.206
		(0.54)	(0.99)	(0.78)
Real Disposable Income Per Capita		-0.012**	0.000	0.002
		(0.00)	(0.00)	(0.00)
Historical Inflation Volatility		-5.772***	-3.847**	-6.482***
		(0.69)	(1.48)	(0.87)
ρ **(Unemployment, Short-term IR)**		23.764**	33.128	28.726**
		(7.87)	(18.53)	(9.79)
Outstanding Covered Bonds to GDP		1.436		5.754***
		(1.24)		(0.80)
Outstanding RMBS to GDP		0.314		2.756***
		(0.82)		(0.50)
Quarter-banking group FE	YES	YES		
Month-country FE	YES	-		
Two-way cluster	-	*country, quarter*	*country, quarter*	*country, quarter*
N	1085	1085	344	344
R^2	0.847	0.779	0.337	0.503
Adjusted R^2	0.733	0.750	0.327	0.492
F-test statistic regressors		-	-	-
degrees of freedom		-	-	-
F-test statistic regressors pure demand		129.047***		53.302***
degrees of freedom		(5,5)		(5,5)
F-test statistic regressors institutional demand		1.957		27.071***
degrees of freedom		(2,5)		(2,5)
F-test statistic fixed effects	7.437***			
degrees of freedom	(464,621)			

* p<0.10, ** p<0.05, *** p<0.01

Table 2.10: Magnitude effects

Magnitude effects of the explanatory variables capturing demand conditions in model 4 of Table 2.9. In the third column the magnitude effect is computed as the product between the estimated coefficient and the standard deviation of the corresponding explanatory variable. In the last column the magnitude effect is computed as the product between the estimated coefficient and the interquartile range of the corresponding explanatory variable.

Variable	**Coefficients**	**Standard deviation**	**Magnitude effect (sd)**	**Interquartile range**	**Magnitude effect (ir)**
Financial Literacy	-5.386**	7.837	-42.213	3.000	-16.159
Indebtedness	0.206	11.097	2.291	12.870	2.657
Real Disposable Income Per Capita	0.002	2215.540	4.386	3260.000	6.453
Historical Inflation Volatility	-6.482***	9.064	-58.758	6.305	-40.872
ρ (Unemployment, Short-term IR)	28.726**	0.491	14.111	0.439	12.608
Outstanding Covered Bonds to GDP	5.754***	5.535	31.854	7.246	41.699
Outstanding RMBS to GDP	2.756***	6.330	17.445	9.652	26.603

Table 2.11: Time variation

Sensitivity of the share of FRMs to the term spread. The term spread is calculated as the difference between the 10-year Interest Rate Swap rate and the 3-month Overnight Index Swap rate. Dependent variable: share of FRMs. Sample: cross-border banking groups. Model: linear. Standard errors: not adjusted. Shorrocks-Shapely decomposition of the R^2 in model 6.

	(1)	(2)	(3)	(4)	(5)	(6)
Austria x tspread				-9.124**		-9.643
				(3.73)		(5.87)
Belgium x tspread				-24.195***		-36.813***
				(2.47)		(4.47)
Germany x tspread				-2.031*		-9.536**
				(1.10)		(4.83)
Spain x tspread				-2.260		-11.414*
				(3.91)		(6.00)
France x tspread				-7.693***		-23.565***
				(1.24)		(4.17)
Italy x tspread				-8.795***		-16.524***
				(1.42)		(4.43)
Luxembourg x tspread				-14.020***		-18.195***
				(2.05)		(2.46)
Slovenia x tspread				-27.161***		-46.651***
				(2.32)		(4.87)
Country FE	YES	-	YES	YES	-	YES
Banking group FE	-	YES	YES	-	YES	YES
Banking group FE x term spread	-	-	-	-	YES	YES
N	1644	1644	1644	1644	1644	1644
R^2	0.580	0.1458	0.6199	0.657	0.206	0.709
Adjusted R^2	0.578	0.1437	0.6173	0.654	0.201	0.705
R^2 country FE						0.303
R^2 country FE x term spread						0.279
R^2 banking group FE						0.054
R^2 banking group FE x term spread						0.073
F-test statistic	322.089***	69.927***	241.939***	207.597***	46.976***	171.713***
degrees of freedom	(7,1636)	(4,1639)	(11,1632)	(15,1628)	(9,1634)	(23,1620)

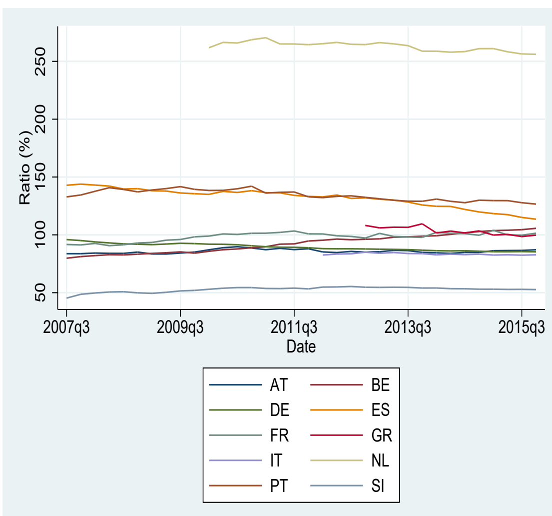
* p<0.10, ** p<0.05, *** p<0.01

2.11 Appendix

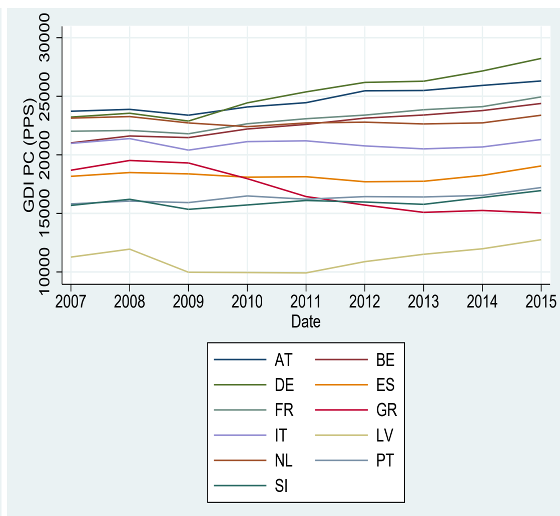
Figure A2.1: Explanatory variables

Explanatory variables: Indebtedness, Real Disposable Income Per Capita, Financial Literacy and Historical Inflation Volatility. Indebtedness is the ratio of total outstanding debt as percentage of gross disposable income provided by the OECD on a quarterly frequency. Data are missing for Latvia and Luxembourg, and partially available for Greece, Italy and the Netherlands. Real Disposable Income Per Capita is the gross disposable income (adjusted for social transfers in kind) of households (and non-profit institutions serving households) expressed in purchasing power standard (PPS) per inhabitant, obtained from Eurostat on an annual basis. Data are missing for Luxembourg. Financial Literacy is measured as the percentage of 3 out of 4 answers correct given by adults interviewed in each country, as results from the S&P Global FinLit Survey. Historical Inflation Volatility is the realized standard deviation of the monthly month-on-month inflation rate during the period 1970-1999. Because of a lack in the available data, Historical Inflation Volatility is computed over the period 1991-1999 for Latvia and 1980-1999 for Slovenia.

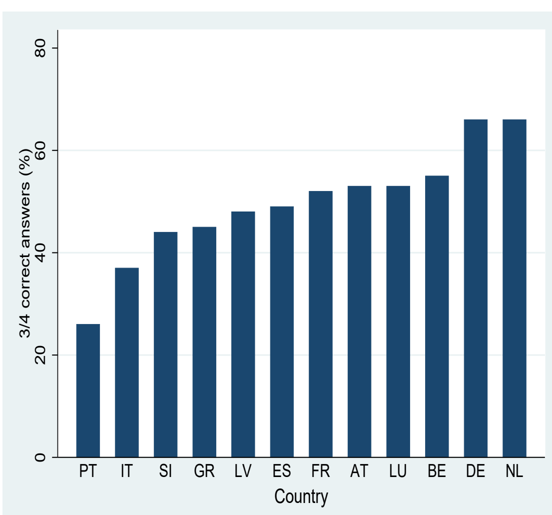
a. Indebtedness



b. Real Disposable Income Per Capita



c. Financial Literacy



d. Historical Inflation Volatility

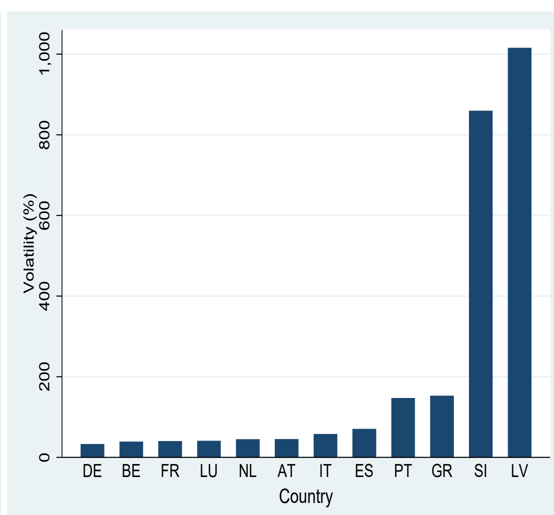
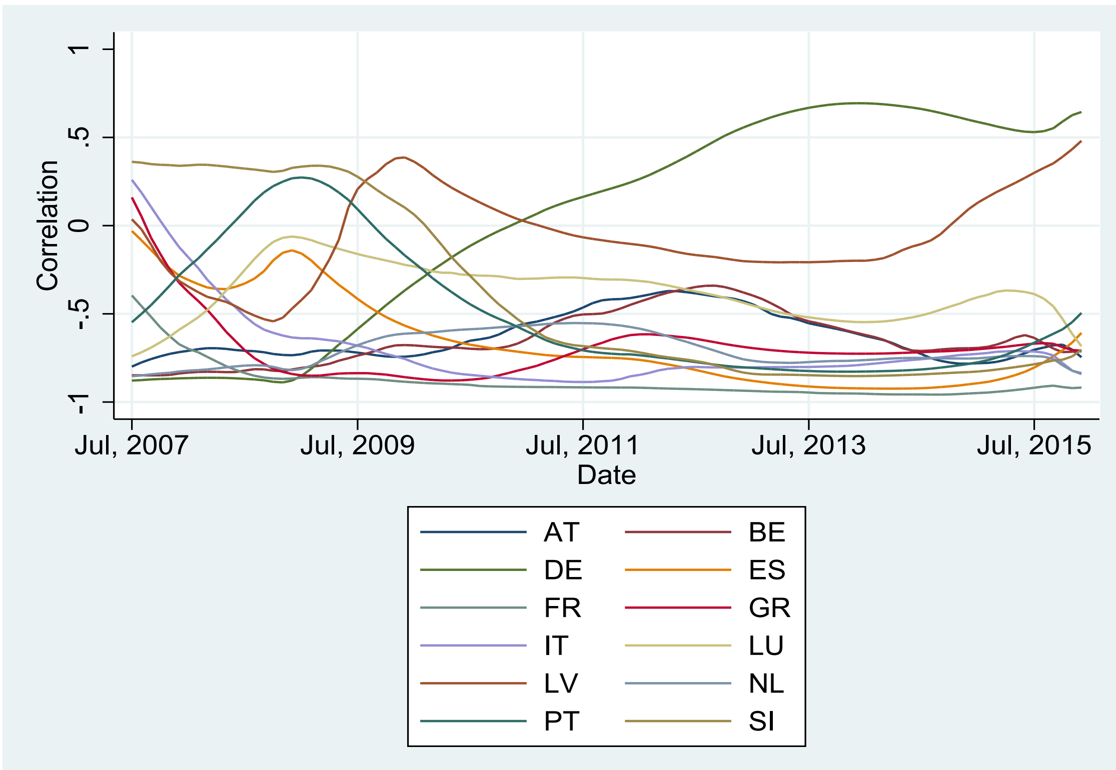


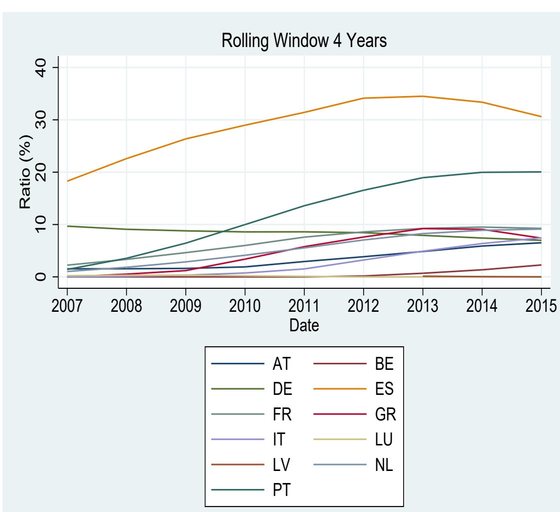
Figure A2.2: Explanatory variables

Explanatory variables: ρ (Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. ρ (Unemployment, Short-term IR) is the realized correlation between the unemployment rate and the short-term interest rate, calculated on a rolling window approach with a window of 7 years. Outstanding Covered Bonds to GDP is the average over the last four years of the amount outstanding of mortgage covered bonds as percentage of GDP. Data are missing for Slovenia. Outstanding RMBS to GDP is the average over the last four quarters of the amount outstanding of RMBS as percentage of GDP. Data are missing for Latvia, Luxembourg and Slovenia and not available for all other countries in 2007.

a. ρ (Unemployment, Short-term IR)



b. Outstanding Covered Bonds to GDP



c. Outstanding RMBS to GDP

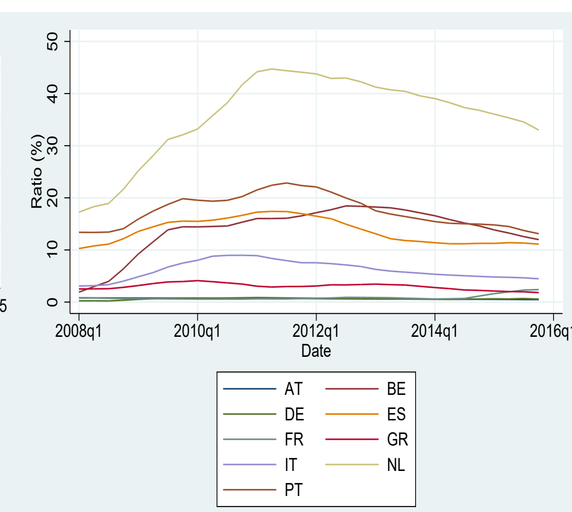


Table A2.1: Fixed effects decomposition with time invariant fixed effects

Fixed effects decomposition of the share of FRMs. Sample: cross-border banking groups. Model: linear. Dependent variable: share of FRMs. Standard errors: not adjusted.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Month FE	YES	-	-	-	YES	YES	YES
Country FE	-	YES	-	YES	YES	-	YES
Banking group FE	-	-	YES	YES	-	YES	YES
N	1644	1644	1644	1644	1644	1644	1644
R^2	0.080	0.580	0.146	0.620	0.669	0.223	0.708
Adjusted R^2	0.019	0.578	0.144	0.617	0.646	0.170	0.687
F-test statistic	1.319**	322.089***	69.927***	241.939***	28.712***	4.194***	33.198***
degrees of freedom	(102,1542)	(8,1636)	(5,1639)	(12,1632)	(109,1535)	(106,1538)	(113,1531)

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table A2.2: Advanced model, one-way clustering

Decomposition of the share of FRMs. Sample: cross-border banking groups. Model: linear. Dependent variable: share of FRMs. Explanatory variables: Financial Literacy, Indebtedness, Real Disposable Income Per Capita, Historical Inflation Volatility, ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. Standard errors: not adjusted in model 1-2, one-way clustered in model 3-14.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Financial Literacy	-0.509	-1.689*	-0.509	-1.689	-0.509	-1.689*	-0.509	-1.689	-0.509	-1.689*	-0.509	-1.689	-0.509	-1.689
	(0.40)	(0.95)	(2.52)	(1.27)	(1.90)	(0.65)	(2.04)	(1.13)	(0.81)	(0.91)	(1.30)	(1.40)	(1.60)	(1.70)
Indebtedness	0.835***	0.602***	0.835	0.602	0.835	0.602	0.835	0.602	0.835***	0.602***	0.835***	0.602**	0.835*	0.602
	(0.12)	(0.15)	(0.92)	(0.88)	(0.43)	(0.51)	(0.56)	(0.52)	(0.14)	(0.16)	(0.24)	(0.27)	(0.37)	(0.38)
Real Disposable Income Per Capita	-0.014***	-0.012***	-0.014***	-0.012**	-0.014**	-0.012***	-0.014**	-0.012**	-0.014***	-0.012***	-0.014***	-0.012***	-0.014***	-0.012**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Historical Inflation Volatility	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***
	(0.34)	(0.51)	(1.63)	(0.91)	(0.87)	(0.40)	(1.28)	(0.52)	(0.47)	(0.44)	(0.76)	(0.67)	(0.93)	(0.73)
ρ **(Unemployment, Short-term IR)**	20.473***	24.170***	20.473	24.170**	20.473	24.170*	20.473	24.170**	20.473***	24.170***	20.473***	24.170***	20.473**	24.170***
	(2.90)	(3.25)	(12.93)	(9.12)	(14.36)	(9.77)	(13.10)	(7.44)	(4.38)	(3.79)	(6.96)	(5.87)	(7.26)	(6.59)
Outstanding Covered Bonds to GDP		1.430**		1.430		1.430		1.430		1.430**		1.430		1.430
		(0.72)		(1.35)		(1.34)		(1.21)		(0.55)		(0.89)		(1.06)
Outstanding RMBS to GDP		0.319		0.319		0.319		0.319		0.319		0.319		0.319
		(0.48)		(1.41)		(0.67)		(0.71)		(0.44)		(0.71)		(0.99)
Month-banking group FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
One-way cluster	-		*bank*	*bank*	*group*	*group*	*country*	*country*	*month*	*month*	*quarter*	*quarter*	*year*	*year*
N	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085
R^2	0.785	0.789	0.785	0.789	0.678	0.684	0.785	0.684	0.785	0.789	0.785	0.789	0.785	0.789
Adjusted R^2	0.677	0.682	0.677	0.682	0.677	0.682	0.677	0.682	0.677	0.682	0.677	0.682	0.677	0.682
F-test statistic regressors	303.309***	222.139***	57.537***	66.989***	-	-	4160.566***	-	400.599***	300.034***	160.953***	124.126***	3094.012***	20598.318***
degrees of freedom	(5,720)	(7,718)	(5,14)	(7,14)	-	-	(5,5)	-	(5,92)	(7,92)	(5,30)	(7,30)	(5,7)	(7,7)

* p<0.10, ** p<0.05, *** p<0.01

Table A2.3: Advanced model, two-way clustering

Decomposition of the share of FRMs. Sample: cross-border banking groups. Model: linear. Dependent variable: share of FRMs. Explanatory variables: Financial Literacy, Indebtedness, Real Disposable Income Per Capita, Historical Inflation Volatility, ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. Standard errors: two-way clustered.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Financial Literacy	-0.509	-1.689	-0.509	-1.689	-0.509	-1.689	-0.509	-1.689	-0.509	-1.689	-0.509	-1.689	-0.509	-1.689	-0.509	-1.689	-0.509	-1.689
	(2.11)	(1.16)	(2.17)	(1.28)	(2.25)	(1.24)	(1.90)	(0.75)	(1.88)	(0.86)	(1.85)	(1.22)	(1.74)	(1.29)	(1.84)	(1.47)	(1.93)	(1.64)
Indebtedness	0.835	0.602	0.835	0.602	0.835	0.602	0.835	0.602	0.835	0.602	0.835	0.602	0.835	0.602	0.835	0.602	0.835	0.602
	(0.75)	(0.71)	(0.75)	(0.71)	(0.75)	(0.74)	(0.43)	(0.51)	(0.42)	(0.51)	(0.39)	(0.50)	(0.47)	(0.53)	(0.48)	(0.55)	(0.50)	(0.56)
Real Disposable Income Per Capita	-0.014***	-0.012***	-0.014***	-0.012***	-0.014***	-0.012**	-0.014**	-0.012***	-0.014**	-0.012***	-0.014**	-0.012**	-0.014***	-0.012**	-0.014***	-0.012**	-0.014***	-0.012*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Historical Inflation Volatility	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***	-5.221***	-5.799***
	(1.35)	(0.77)	(1.38)	(0.79)	(1.38)	(0.77)	(0.86)	(0.43)	(0.84)	(0.47)	(0.74)	(0.54)	(1.08)	(0.59)	(1.14)	(0.68)	(1.15)	(0.66)
ρ(**Unemployment, Short-term IR**)	20.473*	24.170***	20.473*	24.170***	20.473*	24.170**	20.473	24.170*	20.473	24.170*	20.473	24.170*	20.473	24.170**	20.473	24.170**	20.473	24.170**
	(10.73)	(7.72)	(10.94)	(8.03)	(10.26)	(7.71)	(14.28)	(9.81)	(14.13)	(9.82)	(13.39)	(9.59)	(10.93)	(7.78)	(11.23)	(8.18)	(10.51)	(7.95)
Outstanding Covered Bonds to GDP		1.430		1.430		1.430		1.430		1.430		1.430		1.430		1.430		1.430
		(1.11)		(1.09)		(0.78)		(1.32)		(1.29)		(1.28)		(1.25)		(1.31)		(1.30)
Outstanding RMBS to GDP		0.319		0.319		0.319		0.319		0.319		0.319		0.319		0.319		0.319
		(1.17)		(1.18)		(1.16)		(0.70)		(0.73)		(0.89)		(0.78)		(0.88)		(1.02)
Month-banking group FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Two-way cluster	*bank, month*	*bank, month*	*bank, quarter*	*bank, quarter*	*bank, year*	*bank, year*	*group, month*	*group, month*	*group, quarter*	*group, quarter*	*group, year*	*group, year*	*country, month*	*country, month*	*country, quarter*	*country, month*	*country, year*	*country, year*
N	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085	1085
R^2	0.785	0.789	0.785	0.789	0.785	0.789	0.785	0.789	0.785	0.789	0.785	0.789	0.785	0.789	0.785	0.789	0.785	0.789
Adjusted R^2	0.677	0.682	0.677	0.682	0.677	0.682	0.677	0.682	0.677	0.682	0.677	0.682	0.677	0.682	0.677	0.682	0.677	0.682
F-test statistic regressors	77.266***	99.520***	70.149***	167.418***	-	-	-	-	-	-	-	-	226.419***	401.801***	276.015***	-	-	-
degrees of freedom	(5,14)	(7,14)	(5,14)	(7,14)	-	-	-	-	-	-	-	-	(5,5)	(7,5)	(5,5)	-	-	-
F-test statistic regressors pure demand		132.435***		155.015***		71489.080***		3270.95***		385.22***		83.07		233.589***		158.955***		690.636***
degrees of freedom		(5,14)		(5,14)		(5,7)		(4,3)		(4,3)		(4,3)		(5,5)		(5,5)		(5,5)
F-test statistic regressors institutional demand		1.234		1.314		3.111		8.1**		3.64***		2.76***		1.997		1.829		2.139
degrees of freedom		(2,14)		(2,14)		(2,7)		(2,3)		(2,3)		(2,3)		(2,5)		(2,5)		(2,5)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.4: Advanced model, quarter-banking group fixed effects and two-way clustering

Decomposition of the share of FRMs. Sample: cross-border banking groups. Model: linear. Dependent variable: share of FRMs. Explanatory variables: Financial Literacy, Indebtedness, Real Disposable Income Per Capita, Historical Inflation Volatility, ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. Standard errors: two-way clustered.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Financial Literacy	-1.634	-1.634	-1.634	-1.634	-1.634	-1.634	-1.634	-1.634	-1.634
	(1.17)	(1.21)	(1.17)	(0.76)	(0.82)	(1.17)	(1.27)	(1.37)	(1.54)
Indebtedness	0.586	0.586	0.586	0.586	0.586	0.586	0.586	0.586	0.586
	(0.75)	(0.71)	(0.73)	(0.50)	(0.51)	(0.49)	(0.56)	(0.54)	(0.55)
Real Disposable Income Per Capita	-0.012***	-0.012***	-0.012***	-0.012***	-0.012***	-0.012**	-0.012**	-0.012**	-0.012**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Historical Inflation Volatility	-5.772***	-5.772***	-5.772***	-5.772***	-5.772***	-5.772***	-5.772***	-5.772***	-5.772***
	(0.84)	(0.81)	(0.78)	(0.45)	(0.48)	(0.55)	(0.64)	(0.69)	(0.67)
ρ(**Unemployment, Short-term IR**)	23.764***	23.764***	23.764**	23.764*	23.764*	23.764*	23.764**	23.764**	23.764**
	(7.86)	(7.71)	(7.34)	(9.56)	(9.54)	(9.27)	(7.94)	(7.87)	(7.56)
Outstanding Covered Bonds to GDP	1.436	1.436	1.436*	1.436	1.436	1.436	1.436	1.436	1.436
	(1.11)	(1.02)	(0.73)	(1.31)	(1.27)	(1.27)	(1.27)	(1.24)	(1.24)
Outstanding RMBS to GDP	0.314	0.314	0.314	0.314	0.314	0.314	0.314	0.314	0.314
	(1.18)	(1.12)	(1.11)	(0.69)	(0.71)	(0.87)	(0.77)	(0.82)	(0.96)
Quarter-banking group FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Two-way cluster	*bank, month*	*bank, quarter*	*bank, year*	*group, month*	*group, quarter*	*group, year*	*country, month*	*country, quarter*	*country, year*
N	1085	1085	1085	1085	1085	1085	1085	1085	1085
R^2	0.779	0.779	0.779	0.779	0.779	0.779	0.779	0.779	0.779
Adjusted R^2	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750
F-test statistic regressors	89.022***	172.875***	-	-	-	-	-	-	-
degrees of freedom	(7,14)	(7,14)	-	-	-	-	-	-	-
F-test statistic regressors pure demand	118.409***	151.116***	12457.165***	2254.51***	311.74***	80.4***	188.281***	129.047***	355.269***
degrees of freedom	(5,14)	(5,14)	(5,7)	(4,3)	(4,3)	(4,3)	(5,5)	(5,5)	(5,5)
F-test statistic regressors institutional demand	1.168	1.408	3.224	10.49**	4.360	2.96	1.948	1.957	2.148
degrees of freedom	(2,14)	(2,14)	(2,7)	(2,3)	(2,3)	(2,3)	(2,5)	(2,5)	(2,5)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.5: Two-stage model, two-way clustering

Second stage regressions. Sample: cross-border banking groups. Model: linear. Dependent variable: estimated coefficients of month-country fixed effects in model (1) of Table 2.9. Explanatory variables: Financial Literacy, Indebtedness, Real Disposable Income Per Capita, Historical Inflation Volatility, ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. Standard errors: two-way clustered.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Financial Literacy	-2.693	-5.386***	-2.693	-5.386***	-2.693	-5.386**	-2.693	-5.386**	-2.693	-5.386**	-2.693	-5.386**	-2.693	-5.386**	-2.693	-5.386**	-2.693	-5.386**
	(2.16)	(1.42)	(2.15)	(1.60)	(1.89)	(1.56)	(2.18)	(0.96)	(2.18)	(1.12)	(2.17)	(1.04)	(2.28)	(1.53)	(2.26)	(1.72)	(1.99)	(1.62)
Indebtedness	1.558	0.206	1.558	0.206	1.558	0.206	1.558	0.206	1.558	0.206	1.558	0.206	1.558	0.206	1.558	0.206	1.558	0.206
	(0.95)	(0.77)	(0.95)	(0.74)	(0.92)	(0.63)	(1.24)	(0.67)	(1.22)	(0.64)	(1.16)	(0.77)	(0.99)	(0.81)	(0.99)	(0.78)	(0.95)	(0.66)
Real Disposable Income Per Capita	0.000	0.002	0.000	0.002	0.000	0.002	0.000	0.002	0.000	0.002	0.000	0.002	0.000	0.002	0.000	0.002	0.000	0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Historical Inflation Volatility	-3.847**	-6.482***	-3.847**	-6.482***	-3.847**	-6.482***	-3.847	-6.482***	-3.847	-6.482***	-3.847	-6.482***	-3.847*	-6.482***	-3.847**	-6.482***	-3.847**	-6.482***
	(1.45)	(0.73)	(1.42)	(0.80)	(1.18)	(0.59)	(1.70)	(0.72)	(1.77)	(0.78)	(1.98)	(0.40)	(1.52)	(0.82)	(1.48)	(0.87)	(1.22)	(0.60)
ρ(**Unemployment, Short-term IR**)	33.128*	28.726**	33.128*	28.726**	33.128	28.726***	33.128	28.726**	33.128	28.726**	33.128	28.726***	33.128	28.726**	33.128	28.726**	33.128	28.726**
	(17.65)	(9.79)	(17.67)	(9.07)	(18.39)	(7.82)	(16.77)	(5.70)	(16.87)	(5.68)	(16.50)	(1.71)	(18.52)	(10.50)	(18.53)	(9.79)	(19.10)	(8.41)
Outstanding Covered Bonds to GDP		5.754***		5.754***		5.754***		5.754***		5.754***		5.754**		5.754***		5.754***		5.754***
		(0.88)		(0.89)		(0.95)		(0.71)		(0.78)		(1.48)		(0.80)		(0.80)		(0.75)
Outstanding RMBS to GDP		2.756***		2.756***		2.756**		2.756**		2.756**		2.756		2.756***		2.756***		2.756**
		(0.41)		(0.49)		(0.79)		(0.49)		(0.62)		(1.39)		(0.37)		(0.50)		(0.81)
Two-way cluster	*bank, month*	*bank, month*	*bank, quarter*	*bank, quarter*	*bank, year*	*bank, year*	*group, month*	*group, month*	*group, quarter*	*group, quarter*	*group, year*	*group, year*	*country, month*	*country, month*	*country, quarter*	*country, month*	*country, year*	*country, year*
N	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344
R^2	0.337	0.503	0.337	0.503	0.337	0.503	0.337	0.503	0.337	0.503	0.337	0.503	0.337	0.503	0.337	0.503	0.337	0.503
Adjusted R^2	0.327	0.492	0.327	0.492	0.327	0.492	0.327	0.492	0.327	0.492	0.327	0.492	0.327	0.492	0.327	0.492	0.327	0.492
F-test statistic regressors	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
degrees of freedom	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F-test statistic regressors pure demand		41.924***		56.170***		521.38***		107.48***		42.37***		742.40***		44.792***		53.302***		1334.1***
degrees of freedom		(5,8)		(5,8)		(4,7)		(5,3)		(4,3)		(4,3)		(5,5)		(5,5)		(4,5)
F-test statistic regressors institutional demand		25.777***		21.543***		20.475***		44.29***		31.87***		26.64**		34.766***		27.071***		31.334***
degrees of freedom		(2,8)		(2,8)		(2,7)		(2,3)		(2,3)		(2,3)		(2,5)		(2,5)		(2,5)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.6: Two-stage model, year-banking group fixed effects

Two stage regression analysis. First stage regressions including: month-country fixed effects and year-banking group fixed effects (1), all explanatory variables and year-banking group fixed effects (2). In the first stage regressions the dependent variable is the share of FRMs. Second stage regressions of the estimated coefficients of month-country fixed effects in (1) on: pure demand explanatory variables (3) and all explanatory variables (4). Sample: cross-border banking groups. Model: linear. Explanatory variables: Financial Literacy, Indebtedness, Real Disposable Income Per Capita, Historical Inflation Volatility, ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. Standard errors: not adjusted for model (1), two-way clustered by country and quarter for model (2)-(4).

	1ST STAGE		***2ND TAGE***	
	(1)	(2)	(3)	(4)
Financial Literacy		-0.880	-3.772	-6.144***
		(1.30)	(2.06)	(1.38)
Indebtedness		0.566	1.580	0.423
		(0.54)	(0.89)	(0.63)
Real Disposable Income Per Capita		-0.014**	0.000	0.002
		(0.00)	(0.00)	(0.00)
Historical Inflation Volatility		-5.389***	-4.208**	-6.587***
		(0.57)	(1.50)	(0.73)
ρ **(Unemployment, Short-term IR)**		20.945**	41.719*	37.293***
		(6.53)	(18.38)	(8.91)
Outstanding Covered Bonds to GDP		1.053		5.094***
		(1.39)		(0.77)
Outstanding RMBS to GDP		0.026		2.680***
		(0.87)		(0.40)
Year-banking group FE	YES	YES		
Month-country FE	YES	-		
Two-way cluster	-	*country, quarter*	*country, quarter*	*country, quarter*
N	1085	1085	381	381
R^2	0.844	0.758	0.349	0.500
Adjusted R^2	0.748	0.749	0.341	0.490
F-test statistic regressors		-	-	-
degrees of freedom		-	-	-
F-test statistic regressors pure demand		295.46***		68.34***
degrees of freedom		(5,5)		(4,5)
F-test statistic regressors institutional demand		1.940		25.27***
degrees of freedom		(2,5)		(2,5)
F-test statistic fixed effects	8.809***			
degrees of freedom	(413,672)			

* p<0.10, ** p<0.05, *** p<0.01

Table A2.7: Baseline model, Tobit

Fixed effects decomposition of the share of FRMs. Sample: cross-border banking groups. Model: Tobit; lower bound 0 and upper bound 100 in model (1)-(4); lower bound 1 and upper bound 99 in model (5)-(8). Dependent variable: share of FRMs. Standard errors: not adjusted.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Month-country FE	YES	-	YES	YES	YES	-	YES	YES
Month-banking group FE	-	YES	YES	YES	-	YES	YES	YES
Bank FE	-	-	-	YES	-	-	-	YES
N	1644	1644	1644	1644	1644	1644	1644	1644
Pseudo R^2	0.843	0.318	0.908	0.977	0.842	0.318	0.908	0.978
LR test statistic	3047.719***	630.833***	3914.214***	5927.208***	-	631.552***	-	-
degrees of freedom	687	479	1045	1052	-	479	-	-
lower bound	0	0	0	0	1	1	1	1
upper bound	100	100	100	100	99	99	99	99
left censored obs	0	0	0	0	9	9	9	9
right censored obs	0	0	0	0	8	8	8	8

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table A2.8: Fixed effects decomposition with time invariant fixed effects, Tobit

Fixed effects decomposition of the share of FRMs. Sample: all banks. Model: Tobit; lower bound 0 and upper bound 100. Dependent variable: share of FRMs. Standard errors: not adjusted.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Month FE	YES	-	-	-	YES	YES	YES
Country FE	-	YES	-	YES	YES	-	YES
Banking group FE	-	-	YES	YES	-	YES	YES
N	7327	7327	7327	7327	7327	7327	7327
Pseudo R^2	0.026	0.696	0.686	0.778	0.740	0.730	0.825
LR test statistic	191.219***	8740.390***	8515.165***	11051.019***	9714.870***	9436.297***	12483.017***
degrees of freedom	101	11	72	79	112	173	180
lower bound	0	0	0	0	0	0	0
upper bound	100	100	100	100	100	100	100
left censored obs	0	0	0	0	0	0	0
right censored obs	4	4	4	4	4	4	4

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table A2.9: Fixed effects decomposition with time invariant fixed effects, Tobit

Fixed effects decomposition of the share of FRMs. Sample: all banks. Model: Tobit; lower bound 1 and upper bound 99. Dependent variable: share of FRMs. Standard errors: not adjusted.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Month FE	YES	-	-	-	YES	YES	YES
Country FE	-	YES	-	YES	YES	-	YES
Banking group FE	-	-	YES	YES	-	YES	YES
N	7327	7327	7327	7327	7327	7327	7327
Pseudo R^2	0.025	0.697	0.687	0.779	0.741	0.732	0.827
LR test statistic	207.054***	8544.702***	8603.544***	10992.751***	9532.775***	9566.057***	12463.168***
degrees of freedom	101	11	72	79	112	173	180
lower bound	1	1	1	1	1	1	1
upper bound	99	99	99	99	99	99	99
left censored obs	187	187	187	187	187	187	187
right censored obs	214	214	214	214	214	214	214

* p<0.10, ** p<0.05, *** p<0.01

Table A2.10: Advanced model, Tobit

Decomposition of the share of FRMs. Sample: cross-border banking groups. Model: Tobit; lower bound 0 and upper bound 100 in model (1)-(3); lower bound 1 and upper bound 99 in model (4)-(6). Displayed coefficients: marginal effects of the censored variable $\mathbb{E}\left[y\,|x\right]$ at the sample means. Dependent variable: share of FRMs. Explanatory variables: Financial Literacy, Indebtedness, Real Disposable Income Per Capita, Historical Inflation Volatility, ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. Standard errors: one-way clustered by country for model (1)-(2) and model (4)-(5), not adjusted for model (3) and model (6).

	(1)	(2)	(3)	(4)	(5)	(6)
Financial Literacy	-0.502	-1.666		-0.516	-1.677	
	(1.64)	(1.12)		(1.66)	(1.10)	
Indebtedness	0.823*	0.593		0.820*	0.588	
	(0.45)	(0.51)		(0.46)	(0.52)	
Real Disposable Income Per Capita	-0.014***	-0.012***		-0.014***	-0.012***	
	(0.00)	(0.00)		(0.00)	(0.00)	
Historical Inflation Volatility	-5.146***	-5.720***		-5.164***	-5.737***	
	(1.04)	(0.57)		(1.05)	(0.56)	
ρ (**Unemployment, Short-term IR**)	20.180*	23.842***		19.999*	23.665***	
	(10.56)	(7.35)		(10.64)	(7.39)	
Outstanding Covered Bonds to GDP		1.411			1.416	
		(1.20)			(1.18)	
Outstanding RMBS to GDP		0.314			0.309	
		(0.70)			(0.69)	
Month-banking group FE	YES	YES	YES	YES	YES	YES
Month-country FE	-	-	YES	-	-	YES
One-way cluster	*country*	*country*	-	*country*	*country*	-
N	1085	1085	1085	1085	1085	1085
Pseudo R^2	0.787	0.791	0.852	0.787	0.790	0.852
LR test statistic			2075.750***			-
degrees of freedom			605			-
F-test statistic regressors	6263.96***	2054.02***		3.9e+06***	1472.68***	
degrees of freedom	(5,721)	(7,719)		(5,721)	(5,719)	
F-test statistic regressors pure demand		433.93***			435.72***	
degrees of freedom		(5,719)			(5,719)	
F-test statistic regressors institutional demand		2.16			2.20	
degrees of freedom		(2,719)			(2,719)	
lower bound	0	0	0	1	1	1
upper bound	100	100	100	99	99	99
left censored obs	0	0	0	3	3	3
right censored obs	0	0	0	6	6	6

* p<0.10, ** p<0.05, *** p<0.01

Table A2.11: Two-stage model, Tobit

Two stage regression analysis. First stage regressions including: month-country fixed effects and quarter-banking group fixed effects (1), all explanatory variables and quarter-banking group fixed effects (2). In the first stage regressions the dependent variable is the share of FRMs. Model: Tobit, lower bound 0 and upper bound 100 in model (1)-(2). Displayed coefficients: marginal effects of the censored variable $\mathbb{E}\left[y\,|x\right]$ at the sample means. Second stage regressions of the estimated coefficients of month-country fixed effects in (1) on: pure demand explanatory variables (3) and all explanatory variables (4). Sample: cross-border banking groups. Model: linear. Explanatory variables: Financial Literacy, Indebtedness, Real Disposable Income Per Capita, Historical Inflation Volatility, ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. Standard errors: not adjusted for model (1), one-way clustered by country for model (2), two-way clustered by country and quarter for model (3)-(4).

	1ST STAGE			***2ND TAGE***	
	(1)	(2)		(3)	(4)
Financial Literacy		-1.608		-2.693	-5.386**
		(1.08)		(2.26)	(1.72)
Indebtedness		0.576		1.558	0.206
		(0.51)		(0.99)	(0.78)
Real Disposable Income Per Capita		-0.012***		0.000	0.002
		(0.00)		(0.00)	(0.00)
Historical Inflation Volatility		-5.682***		-3.847**	-6.482***
		(0.60)		(1.48)	(0.87)
ρ (**Unemployment, Short-term IR**)		23.390***		33.128	28.726**
		(7.20)		(18.53)	(9.79)
Outstanding Covered Bonds to GDP		1.414			5.754***
		(1.16)			(0.80)
Outstanding RMBS to GDP		0.309			2.756***
		(0.67)			(0.50)
Quarter-banking group FE	YES	YES			
Month-country FE	YES	-			
Clustering	-	*country*		*country, quarter*	*country, quarter*
N	1085	1085	N	344	344
Pseudo R^2	0.847	0.780	R^2	0.337	0.503
LR test statistic	2038.38***		Adjusted R^2	0.327	0.492
degrees of freedom	463				
F-test statistic regressors		193.13***		-	-
degrees of freedom		(5,959)		-	-
F-test statistic regressors pure demand		493.94***			53.30***
degrees of freedom		(5, 959)			(5,5)
F-test statistic regressors institutional demand		2.36*			27.07***
degrees of freedom		(2, 959)			(2,5)
lower bound	0	0			
upper bound	100	100			
left censored obs	0	0			
right censored obs	0	0			

* p<0.10, ** p<0.05, *** p<0.01

Table A2.12: Two-stage model, Tobit

Two stage regression analysis. First stage regressions including: month-country fixed effects and quarter-banking group fixed effects (1), all explanatory variables and quarter-banking group fixed effects (2). In the first stage regressions the dependent variable is the share of FRMs. Model: Tobit, lower bound 1 and upper bound 99 in model (1)-(2). Displayed coefficients: marginal effects of the censored variable $\mathbb{E}\left[y\,|x\right]$ at the sample means. Second stage regression of the estimated coefficients of month-country fixed effects in (1) on: pure demand explanatory variables (3) and all explanatory variables (4). Sample: cross-border banking groups. Model: linear. Explanatory variables: Financial Literacy, Indebtedness, Real Disposable Income Per Capita, Historical Inflation Volatility, ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. Standard errors: not adjusted for model (1), one-way clustered by country for model (2), two-way clustered by country and quarter for model (3)-(4).

	1ST STAGE			***2ND TAGE***	
	(1)	(2)		(3)	(4)
Financial Literacy		-1.599		-2.904	-5.744**
		(1.05)		(2.33)	(1.73)
Indebtedness		0.572		1.506	0.087
		(0.51)		(1.03)	(0.79)
Real Disposable Income Per Capita		-0.012***		0.000	0.002
		(0.00)		(0.00)	(0.00)
Historical Inflation Volatility		-5.690***		-4.079**	-6.857***
		(0.59)		(1.53)	(0.88)
ρ (**Unemployment, Short-term IR**)		23.163***		33.141	28.596**
		(7.20)		(19.26)	(10.05)
Outstanding Covered Bonds to GDP		1.406			6.052***
		(1.13)			(0.84)
Outstanding RMBS to GDP		0.295			2.882***
		(0.65)			(0.51)
Quarter-banking group FE	YES	YES			
Month-country FE	YES	-			
Clustering	-	*country*		*country, quarter*	*country, quarter*
N	1085	1085	N	344	344
Pseudo R^2	0.847	0.780	R^2	0.337	0.509
LR test statistic	-		Adjusted R^2	0.327	0.499
degrees of freedom	-				
F-test statistic regressors		191.75***		-	-
degrees of freedom		(5,959)		-	-
F-test statistic regressors pure demand		511.37***			52.55***
degrees of freedom		(5, 959)			(5, 5)
F-test statistic regressors institutional demand		2.41*			27.65***
degrees of freedom		(2, 959)			(2, 5)
lower bound	1	1			
upper bound	99	99			
left censored obs	3	3			
right censored obs	6	6			

* p<0.10, ** p<0.05, *** p<0.01

Table A2.13: Time variation, Tobit

Sensitivity of the share of FRMs to the term spread. The term spread is calculated as the difference between the 10-year Interest Rate Swap rate and the 3-month Overnight Index Swap rate. Dependent variable: share of FRMs. Sample: cross-border banking groups. Model: Tobit, lower bound 0 and upper bound 100 for model 1-3; Tobit, lower bound 1 and upper bound 99 for model 4-6. Displayed coefficients: marginal effects of the censored variable $\mathbb{E}[y|x]$ at the sample means.

	(1)	(2)	(3)	(4)	(5)	(6)
Austria x tspread	-6.009***		0.249	-6.221***		-0.392
	(2.30)		(3.53)	(2.21)		(3.52)
Belgium x tspread	-21.005***		-24.816***	-20.683***		-24.485***
	(2.19)		(2.49)	(2.18)		(2.49)
Germany x tspread	-1.556*		0.252	-1.519*		0.256
	(0.84)		(1.19)	(0.82)		(1.16)
Spain x tspread	-1.755		-1.272	-2.199		-1.733
	(2.96)		(3.31)	(2.86)		(3.25)
France x tspread	-6.736***		-12.875***	-6.757***		-12.948***
	(1.08)		(1.73)	(1.07)		(1.73)
Italy x tspread	-8.706***		-6.58***	-8.687***		-6.536***
	(1.40)		(1.44)	(1.41)		(1.45)
Luxembourg x tspread	-13.716***		-7.159***	-13.663***		-6.96**
	(2.01)		(2.76)	(2.02)		(2.73)
Slovenia x tspread	-23.464***		-30.941***	-23.157***		-30.493***
	(2.02)		(2.44)	(2.00)		(2.43)
Country FE	YES	-	YES	YES	-	YES
Banking group FE	-	YES	YES	-	YES	YES
Banking group FE x term spread	-	YES	YES	-	YES	YES
N	1644	1644	1644	1644	1644	1644
Pseudo R^2	0.662	0.204	0.715	0.663	0.204	0.716
LR test statistic	1757.598***	378.309***	2030.115***	1752.183***	374.900***	2020.307***
degrees of freedom	15	9	23	15	9	23
lower bound	0	0	0	1	1	1
upper bound	100	100	100	99	99	99
left censored obs	0	0	0	9	9	9
right censored obs	0	0	0	8	8	8

* p<0.10, ** p<0.05, *** p<0.01

Table A2.14: Baseline model, spread

Fixed effects decomposition of the spread between FRMs and ARMs interest rates. Sample: cross-border banking groups. Model: linear. Dependent variable: spread between FRMs and ARMs interest rates. Standard errors: not adjusted.

	(1)	(2)	(3)	(4)
Month-country FE	YES	-	YES	YES
Month-banking group FE	-	YES	YES	YES
Bank FE	-	-	-	YES
N	1642	1642	1642	1642
R^2	0.605	0.378	0.729	0.873
Adjusted R^2	0.322	0.124	0.256	0.646
F-test statistic	2.139***	1.486***	1.540***	3.842***
degrees of freedom	(686,956)	(478,1164)	(1044,598)	(1055,587)

* p<0.10, ** p<0.05, *** p<0.01

Table A2.15: Two-stage, spread

Two-stage regression analysis of the spread between FRMs and ARMs interest rates. First stage regressions including: month-country fixed effects and year-banking group fixed effects (1), all explanatory variables and year-banking group fixed effects (2). In the first stage regressions the dependent variable is the spread between FRMs and ARMs interest rates. Second stage regressions of the estimated coefficients of month-country fixed effects in (1) on: pure demand explanatory variables (3) and all explanatory variables (4). Sample: cross-border banking groups. Model: linear. Explanatory variables: Financial Literacy, Indebtedness, Real Disposable Income Per Capita, Historical Inflation Volatility, ρ(Unemployment, Short-term IR), Outstanding Covered Bonds to GDP and Outstanding RMBS to GDP. Standard errors: not adjusted for model (1), two-way clustered by country and quarter for model (2)-(4).

	1ST STAGE		***2ND TAGE***	
	(1)	(2)	(3)	(4)
Financial Literacy		-0.044	0.041	0.014
		(0.03)	(0.04)	(0.02)
Indebtedness		-0.006	-0.010	-0.007
		(0.01)	(0.02)	(0.02)
Real Disposable Income Per Capita		-0.000	-0.000	-0.000
		(0.00)	(0.00)	(0.00)
Historical Inflation Volatility		0.008	0.042	0.015
		(0.02)	(0.03)	(0.01)
ρ **(Unemployment, Short-term IR)**		0.012	-0.411	-0.183***
		(0.25)	(0.26)	(0.04)
Outstanding Covered Bonds to GDP		0.045***		0.020
		(0.01)		(0.03)
Outstanding RMBS to GDP		-0.011		-0.039***
		(0.01)		(0.01)
Year-banking group FE	YES	YES		
Month-country FE	YES	-		
Two-way cluster	-	*country, quarter*	*country, quarter*	*country, quarter*
N	1085	1085	381	381
R^2	0.616	0.534	0.249	0.348
Adjusted R^2	0.380	0.517	0.239	0.336
F-test statistic regressors		-	-	-
degrees of freedom		-	-	-
F-test statistic regressors pure demand		34.62***		546.84***
degrees of freedom		(5,5)		(4,5)
F-test statistic regressors institutional demand		9.37**		8.95**
degrees of freedom		(2,5)		(2,5)
F-test statistic fixed effects	2.614			
degrees of freedom	(413,672)			

* p<0.10, ** p<0.05, *** p<0.01

Table A2.16: Time variation, spread

Sensitivity of the spread between FRMs and ARMs interest rates to the term spread. The term spread is calculated as the difference between the 10-year Interest Rate Swap rate and the 3-month Overnight Index Swap rate. Dependent variable: spread between FRMs and ARMs interest rates. Sample: cross-border banking groups. Model: linear. Standard errors: not adjusted. Shorrocks-Shapely decomposition of the R^2 in model 6.

	(1)	(2)	(3)	(4)	(5)	(6)
Austria x tspread				0.288*		0.206
				(0.16)		(0.25)
Belgium x tspread				0.454***		0.183
				(0.10)		(0.19)
Germany x tspread				0.308***		0.282
				(0.05)		(0.20)
Spain x tspread				0.353**		0.332
				(0.17)		(0.25)
France x tspread				0.319***		0.174
				(0.05)		(0.18)
Italy x tspread				0.605***		0.486***
				(0.06)		(0.19)
Luxembourg x tspread				0.644***		0.519***
				(0.09)		(0.11)
Slovenia x tspread				1.082***		1.132***
				(0.10)		(0.21)
Country FE	YES	-	YES	YES	-	YES
Banking group FE	-	YES	YES	-	YES	YES
Banking group FE x term spread	-	-	-	-	YES	YES
N	1644	1644	1644	1644	1644	1644
R^2	0.377	0.198	0.456	0.496	0.294	0.581
Adjusted R^2	0.375	0.196	0.452	0.491	0.290	0.575
R^2 country FE						0.171
R^2 country FE x term spread						0.220
R^2 banking group FE						0.076
R^2 banking group FE x term spread						0.114
F-test statistic	141.490***	101.288***	124.139***	106.504***	75.335***	97.653***
degrees of freedom	(7,1634)	(4,1637)	(11,1630)	(15,1626)	(9,1632)	(23,1618)

* p<0.10, ** p<0.05, *** p<0.01

Chapter 3

Credit and Income

Using a unique data set of business loan applications to a bank from individuals who are majority owners of small firms, we study how bank credit origination or denial affects individuals' income. The bank cutoff rule based on the applicants' credit score creates a sharp discontinuity in the decision to originate loans or not. We show that loan origination increases recipients' income five years onward by more than 10% compared to denied applicants. The effect is more pronounced in rural and low-income areas. Our results suggest an important role for banks' credit decisions in affecting the distribution of income.

Delis, M., Fringuellotti, F., Ongena, S., 2019. Credit and Income. Working Paper.

3.1 Introduction

Over past decades, the gap between the rich and the poor has risen in most OECD countries (OECD, 2015), yielding a lively debate on the sources of this development and the proper measures to contain the problem. The role of finance is at the forefront of the relevant academic literature (e.g., Greenwood and Jovanovic, 1990; Galor and Zeira, 1993; Demirguc-Kunt and Levine, 2009; Beck et al., 2010). This study aims to identify and quantify how banks' credit decisions (credit origination or denial) affects applicants' future income. The findings have important implications for the relation between credit and individuals' income, and reflect on how credit origination or constraints affect the distribution of income.

Theoretically, asymmetric information between lenders and borrowers affects credit availability. Because the enforcement of loan contracts is imperfect, lenders often require borrowers to pledge collateral. Lenders also ration credit based on an expected probability of repayment. In general, credit expansion accompanies a relaxation of credit constraints, leading to more financing opportunities for the full spectrum of potential borrowers (including the poor) and a possible tightening of the income distribution (Banerjee and Newman, 1993; Galor and Zeira, 1993).

However, credit-constrained individuals often have less wealth, and their exclusion from credit can foster persistent income growth and inequality. More specifically, financial frictions in the form of informational asymmetry imply an important role for wealth (or capital) endowment in liquidity creation. The endowment represents a fixed cost for credit access. The relatively poor cannot always overcome it, irrespective of the quality of their investment ideas, due to adverse selection and moral hazard in the loan origination process. Thus, returns on capital can lead to high persistence in income growth only for those with substantial wealth (Piketty, 1997; Mookherjee and Ray, 2003; Demirguc-Kunt and Levine, 2009). Further, returns on investment usually increase with the amount of capital wealthier individuals employ, initiating a second-order effect due to economies of scale in larger projects (e.g., Evans and Jovanovic, 1989; Greenwood and Jovanovic, 1990).

The existence of a causal link between access to credit and income inequality presupposes that banks' credit decisions (positive or negative) and the associated access (or lack thereof) to credit have a direct effect on individuals' income. Take, for example, two individuals with approximately the same income and credit quality. One gets a new business loan approved;

the other does not. If loan origination implies an increase in the income of the former relative to the latter, then credit affects the income distribution.

A simple plot between GDP per capita (or the Gini coefficient) and the ratio of private credit to GDP for 150 countries over 1960-2015, shows that income (income inequality) is strongly and positively (negatively) correlated to private credit from banks and other financial institutions over GDP (Figure 3.1). Of course, this relation cannot be interpreted as causal; it is confounded by reverse causality, meaning that income inequality may actually drive credit expansion (Kumhof and Ranciere, 2010; Rajan, 2010) and/or omitted-variable bias due to factors jointly affecting the distribution of income and the degree of financial depth, which are difficult to measure (e.g., the availability of new investment ideas).

[Insert Figure 3.1 here]

Our study provides the first empirical analysis of how access to credit affects individuals' income by comparing the future incomes of accepted applicants to those of rejected applicants. We identify this effect using a unique data set of business loan applications to a single large European bank. Our focus is on loan applications from small and micro enterprises that are majority-owned by individuals who have an exclusive relationship with the bank (i.e., they do not obtain credit from other regulated commercial banks). For these applicants, the bank has information on the business owners' incomes and decides whether to grant the loans based on a credit score cutoff rule. Specifically, each applicant receives a credit score at the time of the loan application. Credit is granted to applicants whose credit scores are above the cutoff, and denied otherwise.

The uniqueness of our data lies in the available information on the majority owners, which encompasses income, wealth, and the credit scores assigned by the bank, as well as other applicant and firm characteristics. Importantly, the exclusivity of the relationship between the bank and the applicant means that most applicants (accepted and rejected) reapply for loans. This in turn means the bank maintains information on applicants' income after the original credit decision.

The availability of credit score and future applicants' income is crucial for our identification strategy because it allows us to exploit the cutoff rule as a source of exogenous variation in the credit decision. Our approach builds on the idea that individuals whose credit scores are around the cutoff are virtually the same in terms of credit quality, yielding a regression discontinuity design (RDD). This implies identification from comparing changes in the income of accepted and denied applicants, who prior to the bank's credit decision have similar credit scores (including similar incomes).

We show that a loan origination increases the recipient's income five years onward by more than 10% compared to denied applicants, regardless of whether we control for application probability. The economic interpretation of this finding is that marginally accepted applicants benefit from an approximately 10% increase in their incomes compared to marginally rejected applicants, thereby significantly widening income inequality between the two groups. This finding is robust to several re-specifications and is not affected by the mix of the control variables. Further, the RDD passes the tests for credit score manipulation, and the control variables are continuous around the cutoff. Overall, our result suggests that bank credit decisions (loan origination or denial) affect individuals' income in a significant way.

We relate our finding to income inequality by calculating inequality measures (Gini coefficients and Theil indices) for the loan applicants around the cutoff. We show that the Gini and Theil indices increase (wider income distribution) for the sample of individuals five years after the credit decision compared to the year of the credit decision. Using the same inequality measures, we also document tighter income distribution among accepted applicants and wider income distribution among rejected applicants. These findings are consistent with the theory of a negative nexus between finance and inequality when there is access to credit (Greenwood and Jovanovic, 1990).

We further examine the heterogeneity of our findings in interesting subsamples reflecting additional aspects of how credit affects income and its distribution. We first document stronger effects in low-income regions compared to high-income regions. This suggests that a bank's credit decision is even more important for an applicant's future income in low-income regions, thus potentially affecting income distribution within and across regions. Second, we use the Great Recession to examine how an economic crisis and associated credit crunch affect the credit-income relation. The identified effect is somewhat stronger during the crisis period, in line with the premise that a credit crunch causes more harm to people with lower credit scores. From an empirical viewpoint, our study relates to the literature that looks broadly at how financial development and/or credit constraints affect income distribution by relying on aggregate (at the country or regional level) measures of inequality (mostly the Gini index) and financial development. This body of literature provides mixed results. Clarke et al. (2006), Beck et al. (2010), Kappel (2010), Hamori and Hashiguchi (2012), Delis et al. (2014), and Naceur and Zhang (2016), for example, document a negative relation between financial development and income inequality, consistent with the idea that credit expansion implies relaxed credit constraints. Denk and Cournède (2015), Jauch and Watzka (2016), and de Haan and Sturm (2017), however, point instead to a positive relation, suggesting that financial development improves access to credit only for the rich. Our paper also relates to several studies on financial development and inequality (for a thorough review, see Demirguc-Kunt and Levine, 2009). We contribute to this literature documenting the effect of credit origination on income and income inequality at the individual, micro level.

Another strand of related recent literature examines how credit constraints affect economic outcomes using data on loan applications (such as ours). Berg (2018), for example, shows that credit denial has stronger negative real effects on low-liquidity firms, which need to increase cash holdings and dispose of other assets in response to a loan rejection. A broader body of literature documents how financial constraints affect the transmission of a credit shock due to changes in monetary policy (Gertler and Gilchrist, 1994; Kashyap and Stein, 2000; Jiménez et al. 2012) or bank stability (Klein et al, 2002; Gan, 2007; Duchin et al., 2010; Cingano et al., 2013; Chodorow-Reich, 2014; Balduzzi et al., 2017; Bentolila et al., 2017; Acharya et al., forthcoming; Popov and Rocholl, forthcoming). We contribute to this literature showing that the effect of credit origination on income is stronger in low-income regions where individuals are more credit constrained.

From a methodological perspective, we use uniquely granular data from a single bank as in Iyer and Puri (2012), and Berg (2018). The detailed information on loan applications that we exploit ensures that we rigorously assess the effect of credit decisions on individual's income at the micro level.

The next section describes the data set and empirical identification, emphasizing the particular RDD. Section 3.3 presents the empirical results regarding how bank credit decisions affect loan applicants' income; it also links these effects to income distribution. Section 3.4 concludes the paper.

3.2 Data and Empirical Identification

3.2.1 Loan Applications

We use a unique sample of loan applications to a single large European bank. The bank provides credit to a wide array of small and large firms, as well as to consumers, households, and the public sector both domestically and abroad. Our sample is limited to loan applications from individuals, firms and administrations that are located in the country where the bank is headquartered. We use only loan applications from small and micro enterprises that are majority-owned by specific individuals, for which the bank has important information for our analysis.[1] Specifically, we have information on whether the loan is originated or denied, as well as loan characteristics, firm characteristics, and applicant characteristics. For originated loans, loan characteristics include the amount, maturity, collateral, and other features (covenants, performance-pricing provisions). Firm characteristics encompass several accounting variables, such as assets and sales, profits, leverage, etc.

What makes this data unique is information on the applicant (the firm's majority owner). The applicant characteristics include income, assets (wealth), gender, education, relationship with the bank (an exclusive relationship or not), and the credit score assigned by the bank. For two reasons, we focus on loan applications from individuals who have exclusive relationships with their banks. First, the bank has income information for these applicants for several years before and after the loan origination. Second, these applicants are generally unable to obtain credit from another bank, especially if their applications are denied; moreover, they cannot access capital markets due the firm's small size.[2]

Each applicant is given a credit score at the time of the application, and this score is the decisive factor in loan origination. For comparative purposes, we normalize the credit score to be around the cutoff value of 0. The bank originates the loan if the credit score is higher than 0 and denies the loan otherwise. For very few applications (72 cases), this criterion does not hold. These exceptions are possibly due to data-entry mistakes and thus we disregard them in our analysis. We explicitly define the credit score along with all the

[1] Using the European Commission's definition, a small enterprise has total assets less than €10 million; a micro enterprise less than €2 million in assets.

[2] We have information about this exclusivity from the bank. However, the firms can receive credit (obviously at higher rates) in the shadow-banking sector.

variables used in our empirical analysis in Table 3.1 and provide summary statistics in Table 3.2.

[Insert Table 3.1 & Table 3.2 here]

Using this information, we generate a balanced panel data set, where applicants are the cross-sectional unit of the panel and years 2002-2016 are the time unit. For each applicant, we know his/her income and wealth over the full sample period, as well as for at least five years before and after the loan application. This means that the individuals in our sample do not necessarily apply for loans in some years. This sample also includes information for the rest of the applicant and firm characteristics defined in Table 3.1. This stringent cleansing process yields 234,420 observations corresponding to 15,628 individual applicants over 2002-2016.[3] In this panel, there are 61,863 loan applications (the sample in the majority of our empirical tests). We report summary statistics for the variables in Table 3.2.

The mean future income (respectively, in one year, three years, and five years) tends to rise over time for loan applicants. Banks accept (or partially accept) approximately 87% of loan applications and reject 13%. This rejection rate is a bit higher than the rejection rates reported in the European Commission/European Central Bank Survey on access to finance for enterprises (SAFE).[4] The reason is that some missing observations on variables in our empirical analysis correspond to individuals with strong bank ties (i.e., individuals for whom the bank already has information) who are usually not rejected. If anything, this biases our results in favor of denied applicants. However, our identification approach, based on individuals around the credit score cutoff, should mitigate such concern. After its transformation, the mean credit score is positive and equal to approximately 0.1. Average loan duration is roughly three years.

Summary statistics for our control variables show that the mean applicant has tertiary education and total wealth of €187,200 (see Table 3.2). The mean firm size (total assets) is €369,500, and mean firm leverage is 20.7%, which is comparable to European averages (e.g., Carvalho, 2017). Overall, the summary statistics show that our data set is consistent with the mean value of our variables at the European level.

[3]The actual number of loan applications from small and micro enterprises, including business-loan applications from individuals who have nonexclusive relationships with the bank, as well as those from applicants for which we lack dynamic income information, is 513,525.

[4]See, https://ec.europa.eu/growth/content/survey-access-finance-enterprises-safe-was-published-today_ga.

Using data from a single entity is not an unusual practice when the research question is detailed (Adams et al., 2009; Iyer and Puri, 2012; Berg, 2018). In our case, we take advantage of granular application-level data for one bank to document how the decision to grant or deny credit affects individuals' income. Also, the bank that we look at is a major financial institution operating on a national scale. This ensures that the bank is representative enough for the banking system, so that we can reasonably generalize the results of our study.

3.2.2 Empirical Identification

Three important features of our data set are the availability of information about (i) originated and denied loans, (ii) the exclusivity of the relationship between loan applicants and banks, and (iii) applicants' income before and after the loan application. Based on these features, a standard identification method would compare the incomes of approved applicants (the treated group) with the incomes of rejected applicants (the control group) before and after the loan decision. Unfortunately, the treatment here is endogenous to several factors behind the bank's decision to grant the loan, making a differences-in-differences exercise far from optimal.

The fourth and most important feature of our data set for identification purposes is the availability of information on credit scores and the perfect correlation of the scores above the cutoff with loan origination.[5] This implies a sharp discontinuity in treatment as a function of credit score.[6] Therefore, we rely on a sharp RDD using credit score as the assignment (also referred to as "the running" or "the forcing") variable (Imbens and Lemieux, 2008; Lee and Lemieux, 2010).

Assuming that the relation between access to credit and income is linear, the simplest form of the RDD is:

$$y_{i,t+n} = a_0 + a_1 D_{it} + a_2 \left(x_{it} - \bar{x}\right) + u_{it} \tag{3.1}$$

In equation (3.1), y is applicant's i income in the n^{th} year ahead of the loan application, which takes place in year t. D is a binary variable that equals 1 if the credit score is above

[5] This is after dropping the 72 exceptions due to data entry errors.

[6] Berg (2018) exploits a similar type of discontinuity to investigate how loan rejection affects firms' cash holdings.

the cutoff and zero otherwise, which determines whether the loan is granted. Thus, a_1 is the treatment effect. Also, $(x_{it} - \bar{x})$ is the distance between the cutoff and applicant i's credit score given at the time of the loan application.

The distribution of applicant's income depicted in Figure 3.2 exhibits a regular shape. The main assumption for the validity of this model, similar to any other RDD, is that applicants cannot precisely manipulate their credit scores. If applicants, even while having some influence, are unable to manipulate their credit scores precisely, the variation in treatment around the cutoff provides a randomized experiment. The lack of precise manipulation is the most compelling requirement of the RDD vis-à-vis other identification methods, such as differences-in-differences or instrumental variables (Lee and Lemieux, 2010).

[Insert Figure 3.2 here]

Theoretically, precise manipulation is unlikely, as loans officers' prudent behavior should prevent applicants from having exact information on their credit scores. We demonstrate, through a specific statistical test, that this is also unlikely from an econometric viewpoint. Specifically, we test for manipulation of the assignment variable around the cutoff. Self-selection or nonrandom sorting of applicants would entail a discontinuous change in the distribution of the credit score. Figure 3.2 shows that the probability density of the credit score does not jump around the cutoff. In line with the graphical evidence, the formal test of Cattaneo et al. (2018) confirms there is no statistical evidence of manipulation of the forcing variable (see Table 3.3 and Figure 3.3).

[Insert Table 3.3 & Figure 3.3 here]

3.3 Empirical Results

3.3.1 Parametric Model

We first consider estimating equation (3.1) with a parametric model (OLS). We use clustered standard errors at the individual level to ensure robust inference. To allow for a differential effect on the two sides of the cutoff, we include the interaction $D_{it}(x_{it} - \bar{x})$, so that equation (3.1) becomes:

$$y_{i,t+n} = a_0 + a_1 D_{it} + a_2 (x_{it} - \bar{x}) + a_3 D_{it} (x_{it} - \bar{x}) + u_{it} \tag{3.2}$$

The coefficient of interest is a_1, which is the coefficient of the acceptance dummy *Granted*, which captures the treatment effect.

Table 3.4 reports the results. Specifications 1-3 use as a dependent variable the applicants' income one year ahead, three years ahead, and five years ahead of the loan application. We find a positive and statistically significant coefficient on *Granted* in all three specifications. The magnitude of this coefficient suggests a 5.1% increase in the incomes of approved applicants one year ahead of loan origination (column 1), a 7.3% increase three years ahead (column 2), and a 7% increase five years ahead (column 3). Also, the coefficient of the interaction between *Granted* and *Credit Score* is negative and statistically significant three and five years after loan origination, confirming our prior differential effect on the two sides of the cutoff.

[Insert Table 3.4 here]

In specifications 4-6, we introduce the set of loan, firm, and applicant controls variables. Loan controls include the requested amount (*Loan amount*) and loan maturity (*Maturity*). Firm variables include total assets (*Firm size*) and leverage ratio (*Leverage*). Applicant controls include degree of education (*Education*) and income one year before the application (*Income* $t-1$). We provide thorough definitions for these variables in Table 3.1.

Indeed, the results are similar to those in the first three columns and, if anything, slightly strengthen. Being approved for a loan implies an increase in applicant income by 5.4% one year after of the loan decision (column 4), by 7.5% three years after (column 5), and by 7.2% five years after (column 6). Looking at the covariates, most are not statistically significant. This is not surprising, as many of them concur in determining the credit score. Nevertheless, we find a positive and statistically significant coefficient for *Income* $t-1$, suggesting persistence in the outcome variable. Leverage has a positive and significant coefficient, but it is largely collinear with the credit score.[7] We also find a positive coefficient on *Maturity*, although it is significant only in column 4. These results

[7] Our analysis focuses on firms able to raise external funds only by borrowing from the bank under study. In our specifications, we control for the leverage ratio observed in the year of the loan decision. The cutoff rule implies that applicants whose credit scores are above the cutoff are approved for a loan. As a consequence, leverage ratios increase in the year of the loan origination (see Figure 3.5). This explains why our covariate is to a large extent collinear with the credit score.

remain unchanged if we add industry, loan type, and year fixed effects to our specifications (results in Table A3.1 of the Appendix).

3.3.2 Local Linear Regression

The linear model identifies the treatment effect placing equal weight on all information available in the sample. This suggests a potential bias, as it treats observations far from the cutoff in the same way as observations close to the cutoff, and the treatment effect is estimated using two groups of individuals that might not be comparable. To handle this issue, we use a local linear regression (for a general description, see Imbens and Lemieux, 2008, and Calonico et al., 2014). The main advantage of this approach is the assignment of higher weights as we move closer to the cutoff (using a kernel smoother). We determine the optimal bandwidth using the approach in Calonico et al. (2014), and for efficient estimation we mainly base our inference on the local-quadratic bias-correction in Calonico et al. (2018).

Table 3.5 reports the estimates of the average treatment effect for our set of local linear regressions.[8] For each specification, we report the conventional RD estimates with conventional variance estimator (*Conventional*), the bias-corrected RD estimates with conventional variance estimator (*Bias-corrected*), and the bias-corrected RD estimates with robust variance estimator (*Robust*).

Regardless of whether we include (in columns 1-3) or do not include (in columns 4-6) the set of controls, we find that granting a loan has a positive and significant effect on an applicant's future income. Relying on *Robust* estimates for inference, we find an income increase of approximately 6% among approved applicants one year or three years after of the loan origination, and an increase of approximately 11% five years ahead.

Overall, the estimates of the treatment effect are comparable to those in the corresponding regressions of Table 3.4. The magnitudes of the effect are somewhat higher than those of the OLS regressions, especially considering the effect five years ahead. Given the small discrepancy in the results between the parametric and nonparametric RDD and the advantages of the nonparametric RDD highlighted in the literature, we consider this method as our benchmark and we use it in most of our sensitivity tests (unless not applicable).

[Insert Table 3.5 here]

[8] The average treatment effect here is the counterpart of the coefficient of the acceptance dummy in equation (3.2). It is nonparametrically identified as $\tau_{RD} = \lim_{x \to \bar{x}^+} \mathbb{E}[y_{it} \,|x_{it} = x] - \lim_{x \to \bar{x}^-} \mathbb{E}[y_{it} \,|x_{it} = x]$.

An additional merit of the nonparametric RDD is the graphical inspection of the relation between access to credit and income that takes into account any potential nonlinearity. Figure 3.4 depicts applicants' income five years after the loan decision against the credit score (the figure is from column 3 of Table 3.5 and the effective observations used by the local linear regression). The figure shows a clear upward shift in applicants' income around the cutoff. This indicates that the treatment (loan origination) entails a sharp discontinuity in the outcome variable (income), corroborating our methodological approach.

[Insert Figure 3.4 here]

On the use of control variables, a key assumption of the RDD is that the expectation of the outcome variable conditional on the assignment variable is continuous. This requires that the relation between the covariates and the credit score is smooth around the cutoff. A graphical inspection confirms that this condition is fulfilled (Figure 3.5). This means that our baseline model in equation (3.2) is well specified and, using the controls, will not significantly affect our main result.

[Insert Figure 3.5 here]

Despite the advantage of focusing on observations close to the cutoff, the nonparametric approach does not necessarily represent the ideal functional form of the RDD. In light of that, Lee and Lemieux (2010) suggest relying on different bandwidth-selection methods to test if the results are stable across different specifications. Table A3.2 of the Appendix shows that the results remain unchanged when using the mean-squared error (MSE) or the common coverage error (CER) bandwidth selector. Also, Figure 3.6 shows that the significance of Conventional in model (3) is robust to different windows around the cutoff where (small-sample) inference is conducted.[9]

[Insert Figure 3.6 here]

Overall, our analysis shows that credit decisions have real effects on income. Consider two applicants: the first has a credit score slightly above the cutoff; the second has a credit score slightly below the cutoff. At the time of the loan application, the credit quality of these two individuals is virtually the same. However, the cutoff rule implies that credit

[9] Inference in Table 3.5 is based, instead, on large-sample approximations (Calonico et al., 2014).

is granted only to the former. The increase in income experienced after loan origination documents a causal link between access to credit and income.

3.3.3 Robustness Tests

In principle, wealthier individuals should be able to maintain higher incomes over time through higher investment. Accordingly, part of the macro inequality literature highlights the role of initial GDP per capita and suggests controlling for some sort of historical (or initial) wealth conditions when estimating models of inequality (e.g., Li et al., 1998). To this end, we use individual wealth in the first year before the loan application in which this information is available (*Initial wealth*; see Table 3.1).

As with the rest of the control variables, we show in Figure A3.1 that *Initial wealth* is continuous around the cutoff. Of course, adding this variable to our covariates entails a substantial drop in the number of observations in the sample. This is the reason we leave this exercise as a robustness test. The nonparametric results in Table 3.6 show that including initial wealth does not yield significantly different results. If anything, the treatment effect is slightly stronger, with the only exception of the three-year horizon from the loan decision. We obtain similar patterns when using the parametric RDD (Table A3.2 of the Appendix).

[Insert Table 3.6 here]

So far, our framework does not explicitly model the probability to apply for a loan in a specific year. Given that our sample is a balanced panel of bank customers with exclusive relationships and these customers appear in the panel irrespective of whether they apply for a loan in a given year, we can model the probability of receiving a loan application, and examine its effect in our baseline model. Econometrically, this implies limiting a form of selection bias in the estimation of the treatment effect.

Specifically, we use a parametric two-stage selection model as in e.g. Heckman (1976), Dass and Massa (2011), and Jiménez et al. (2014). In the first stage, we estimate the probability that a bank customer applies for a loan in a specific year (probit model). The right-hand side variables in the first stage are those in columns 4-6 of Tables 3.4 and 3.5, excluding the credit score (which is unknown to the applicant) and including *Gender*.[10] In

[10]We find that *Gender* is significantly correlated with the probability to apply for a loan but does not explain income in the baseline specifications.

the second stage, we run a similar regression to the one implied in equation (3.2), in which we use the predicted instantaneous probability of applying for a loan (hazard rate) from the first stage as an additional control variable.

Table 3.7 reports the estimation results. The first-stage results show that income and wealth positively and strongly affect the hazard rate of a loan application, in line with the premise that wealthier individuals are more likely to apply for credit. The same holds for larger and more leveraged firms. Interestingly, we also find that male applicants are 1.6% more likely to apply for credit than female applicants are. The second-stage results are fully in line with Table 3.4, even though the probability of loan application enters with a highly significant and positive coefficient.

To account for selection of loan applicants, we prefer the standard parametric model because it is standard in the applied economics/finance literature, whereas the nonparametric models are quite rare in this respect. However, we do experiment with a semiparametric model, where we save the parametric first-stage prediction and include it in the nonparametric second stage. Again, the results, reported in Table 3.8, are consistent with those of Table 3.5.

[Insert Table 3.7 & 3.8 here]

3.3.4 Reflection on Income Inequality

A natural implication of our key finding is that income distribution changes. Specifically, we expect that a bank's credit decision increases income inequality between groups of individuals who have similar characteristics (individuals around the cutoff) but receive different credit decisions (accept vs. reject). It is difficult to extend this implication to the full array of income distribution, because most people (and certainly the rich) are granted loans. However, we can construct inequality measures around the cutoff for individual income at the time of loan application (t) and five years ahead ($t + 5$). As our sample around the cutoff, we use individuals with credit scores less than the absolute value of 0.1.[11]

Panel A of Table 3.9 reports the results for the Gini coefficient and the Theil index. Both the indices increase from time t to time $t+5$, reflecting higher income inequality. The effect is economically large and equivalent to that identified in Table 3.5. Specifically, the Gini

[11] Alternatively, we use the effective observations left and right of the cutoff produced by the local linear regression in column 6 of Table 3.5. The results are very similar.

coefficient increases by approximately 9% and the Theil index increases by approximately 10%, indicating considerably higher income inequality after the bank credit decisions for the sample of individuals close to the cutoff.

[Insert Table 3.9 here]

In Panel B of Table 3.9, we construct equivalent Gini and Theil indices for accepted and rejected applicants. The indices show that for accepted applicants, the Gini and Theil indices are significantly lower, whereas for the rejected applicants they are higher. This is consistent with the premise that positive credit decisions allow individuals close to the cutoff to increase their incomes, thereby tightening the income distribution among accepted individuals. In contrast, negative credit decisions are consistent with widening income distribution among rejected individuals, who are the relatively poor.

We conduct two more tests to reflect how credit decisions affect income distribution. The first concerns the role of applicant location based on regional income, distinguishing between low-income regions and high-income regions. In Table 3.10, we replicate the analysis in columns 4-6 of Table 3.5, separating our full sample into low-income and high-income regions based on median income. We expect that the income elasticity to credit decisions is higher in low-income regions, where credit constraints should also be relatively higher.[12]

[Insert Table 3.10 here]

As a final test, we consider the role of the Great Recession. During this period, Europe experienced sharp losses in household wealth and aggregate demand, substantial contraction of credit, and increased unemployment (e.g., IMF, 2009; ECB, 2016). In such context, entrepreneurs face riskier investment opportunities and lower profits. This yields increased dependence on bank credit, even for business survival and especially for small firms. If this prevails, loan origination has a stronger effect on applicant incomes during the crisis period, and a negative credit decision widens the income distribution.

To examine the role of the crisis in our results, we split the sample into the 2000-2008 and the 2009-2016 periods. We leave 2008 in the pre-crisis period because credit from banks in

[12] In our sample, the mean value of *Granted* in high-income regions is 0.880; it is 0.853 in the low-income regions.

European countries was still rising that year. Similarly, we include the full period after the crisis because credit from banks to the private sector over GDP decreased in 2009-2016.[13]

Table 3.11 reports the results from the two samples. We find that three to five years after a bank's credit decision, access to credit has a stronger effect on applicant incomes during a crisis than in normal times. In particular, we find that approved applicants' incomes rose 9.6% five years ahead of the loan origination during 2000-2008 (column 6), versus 10.5% in the crisis and post-crisis periods (column 3). We conclude that, in the medium to long run, a loan origination has a stronger effect on applicant incomes during periods of higher credit constraints than in normal times.

[Insert Table 3.11 here]

3.4 Conclusion

Credit constraints potentially hinder income growth opportunities, especially for those with low incomes and a lack of collateral. Using data from business loan applications to a single large European bank, we study and quantify how a bank's credit decisions (acceptance or rejection) affect individuals' future incomes. We look at loan applications from small and micro enterprises that are majority-owned by individuals for which we have detailed information on past and future income, the credit score assigned by the bank, and the exclusivity of relationship lending (among many other applicant and firm characteristics).

Our identification strategy comprises a regression discontinuity design, exploiting exogenous variation in the credit decision from the cutoff rule on the basis of credit score. Essentially, with this strategy we compare individuals with credit scores (and thus very similar characteristics guiding the credit decision) around the cutoff. We show that access to credit has a positive effect on individual income. Specifically, the income of accepted applicants is approximately 4% higher than the income of denied applicants one to three years ahead of the loan decision; this jumps to 10% five years ahead. This finding is robust to several re-specifications and robustness tests.

We also explore how income distribution changes with bank credit decisions. We first show that the Gini and Theil indices increase for individuals around the cutoff, reflecting increased income inequality within that sample. We also show that credit decisions have

[13] See https://data.worldbank.org/indicator/FS.AST.PRVT.GD.ZS?locations=XC.

a somewhat stronger effect on applicants' future incomes in low-income regions (vs. high-income regions) and during the crisis and post-crisis period (vs. the pre-crisis period). These results highlight the heterogeneous effects credit availability has on applicants' future incomes due to increased credit constraints. They also highlight differential effects on income distribution.

Our findings have two key and interrelated economic implications. First, credit decisions strongly affect applicants' future income and its subsequent dynamics, altering lifetime income expectations and potentially applicants' economic decisions. Second, credit decisions exert substantial effects on income inequality between individuals who prior to the credit decision have similar credit scores. Importantly, these effects are more potent for applicants in low-income regions and during crisis and post-crisis periods.

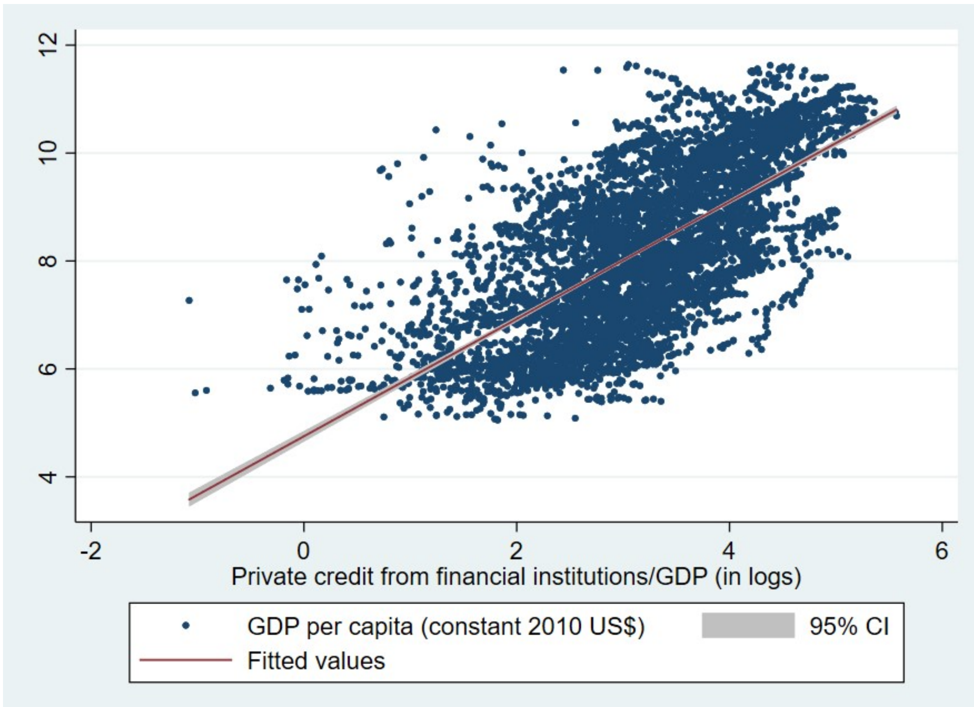
Our findings suggest that an otherwise efficient credit decision affects income distribution and thus supports policy interventions aimed at increasing credit access. Relevant actions are those of the European Bank for Reconstruction and Development (EBRD) and the European Investment Bank (EIB), which selectively target credit-constrained individuals with good investment ideas.

Our findings also open up a discussion on whether central banks (via specialized institutions such as the EBRD and EIB) could direct a small part of the money-creation process to good investment ideas from loan applicants whom the banking system rejected due to a lack of credit history or collateral. We leave the thorough examination of the effects of these policies to future research.

3.5 Figures

Figure 3.1: Income and income inequality against credit

The first graph depicts GDP per capita (in constant 2010 US$) against the ratio of private credit to GDP (x-axis). The second graph depicts the Gini index against the ratio of private credit to GDP (x-axis). We report individual values, as well as fitted values using a linear regression model. The estimated slopes of the linear regressions are 1.087 and -0.077, respectively, and are statistically significant at the 1% level. Data on the Gini index are from the Standardized World Income Inequality Database (SWIID); data on credit and GDP per capita are from the World Development Indicators.



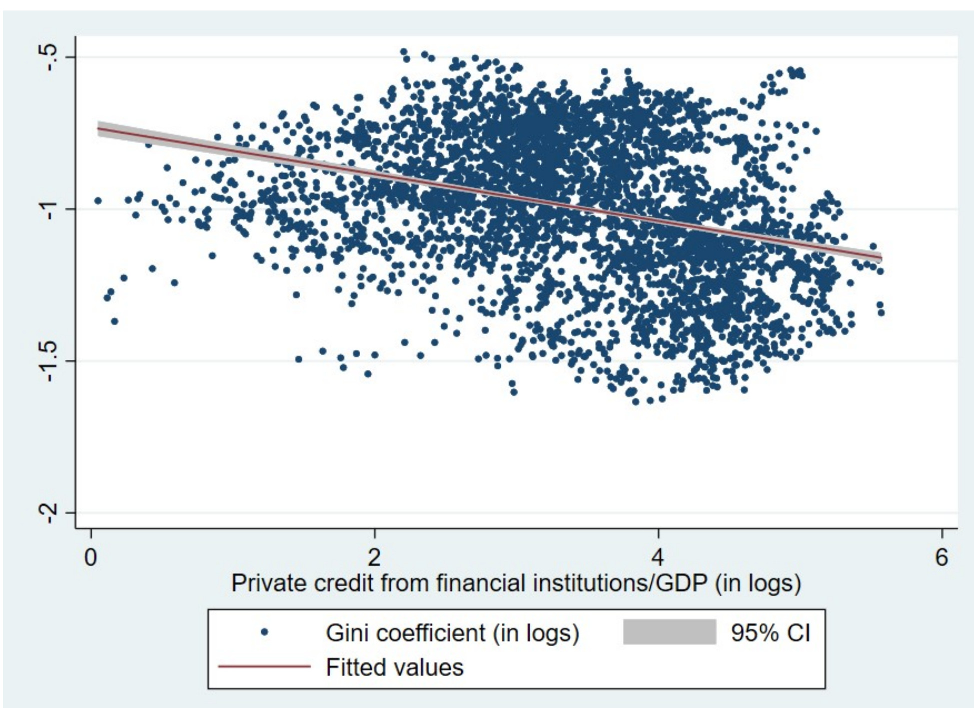
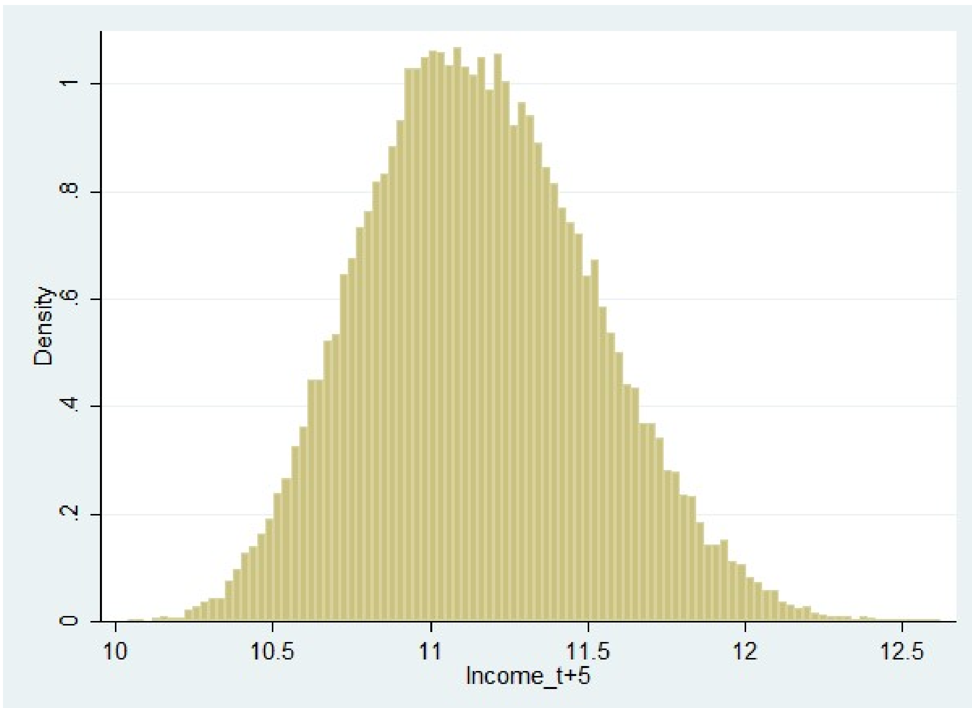


Figure 3.2: Densities of outcome and assignment variables

The figures report the probability densities for the outcome variable Income t+5 (top) and the assignment variable Credit score (bottom).



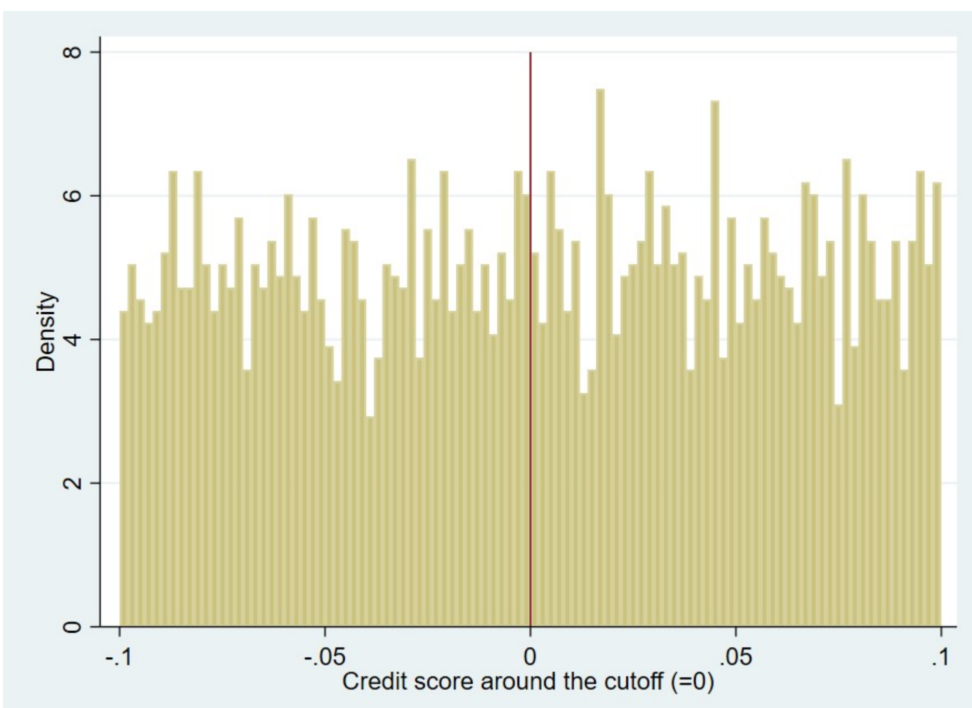


Figure 3.3: Manipulation test

The figure reports results from the manipulation testing procedure using the local polynomial density estimator proposed by Cattaneo et al. (2018). To perform this test, we rely on the local quadratic estimator with cubic bias-correction and triangular kernel.

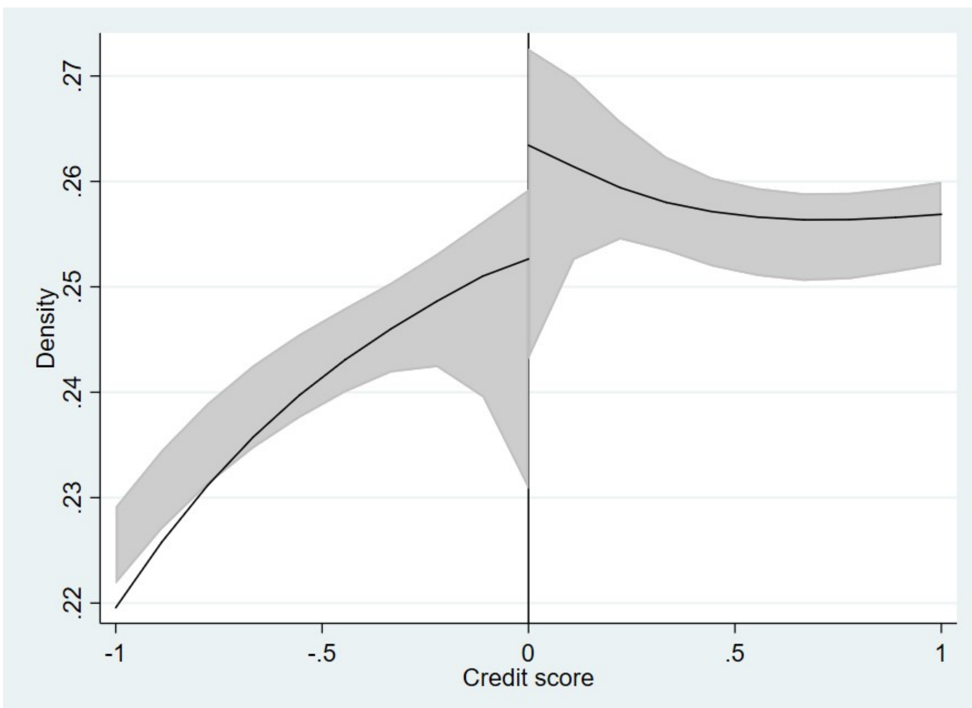


Figure 3.4: Applicants' income around the cutoff

The figure depicts applicants' *Income* five years ahead the loan decision (y-axis) against the *Credit score* (x-axis). The result is obtained from the non-parametric RDD. The continuous line represents a fourth order polynomial fit used to approximate the conditional mean of applicants' income below and above the cutoff.

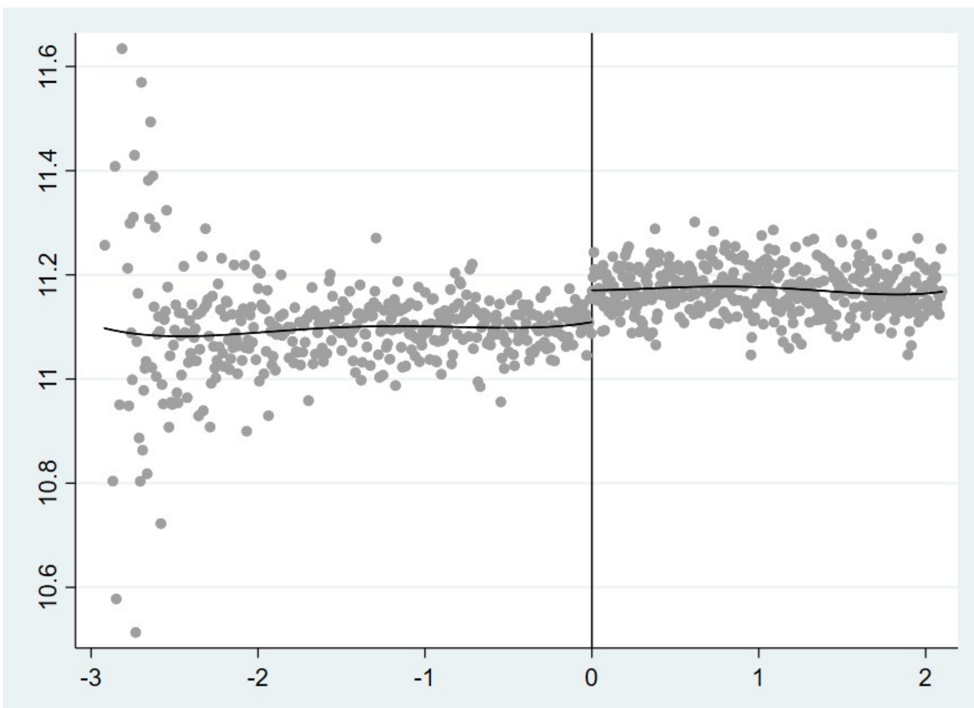
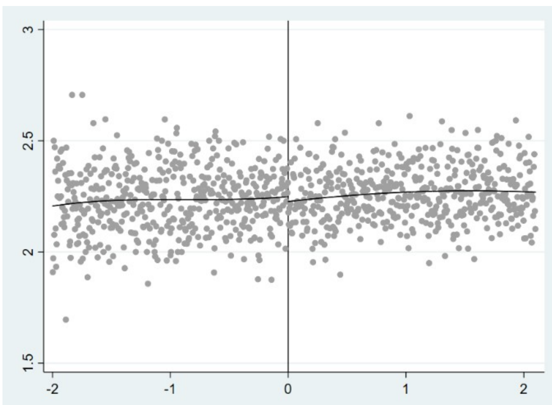


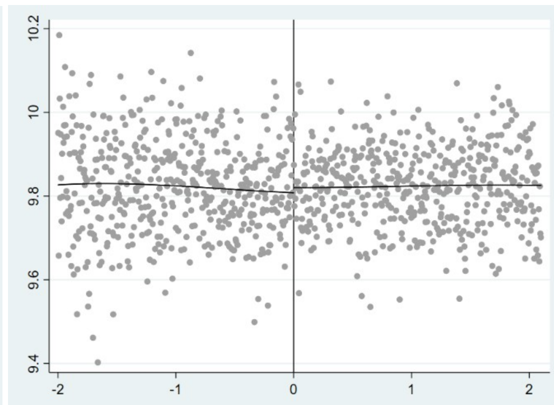
Figure 3.5: Covariates around the cutoff

The figure reports a plot for each control variable against the Credit score. The covariates include Education, Firm size, Firm leverage, Loan amount and Maturity. The continuous line represents a fourth order polynomial fit used to approximate the conditional mean of each covariate below and above the cutoff.

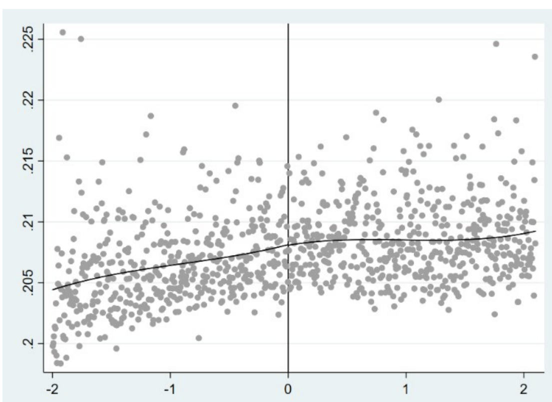
a. Education (y-axis) against Credit score (x-axis)



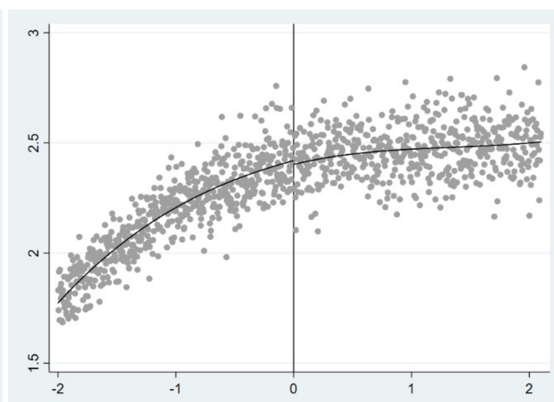
b. Firm size (y-axis) against Credit score (x-axis)



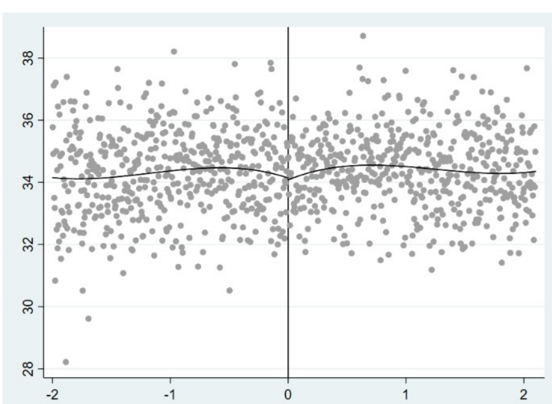
c. Firm leverage (y-axis) against Credit score (x-axis)



d. Loan amount (y-axis) against Credit score (x-axis)



e. Maturity (y-axis) against Credit score (x-axis)



f. Income t-1 (y-axis) against Credit score (x-axis)

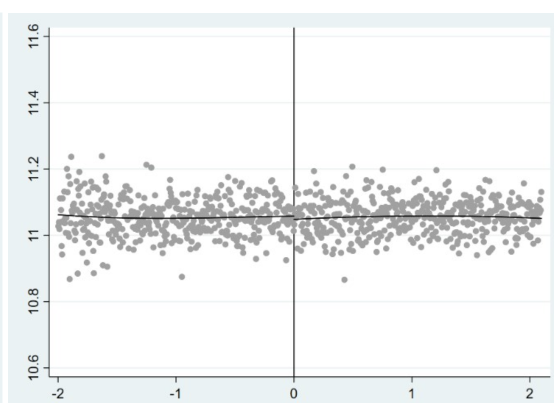
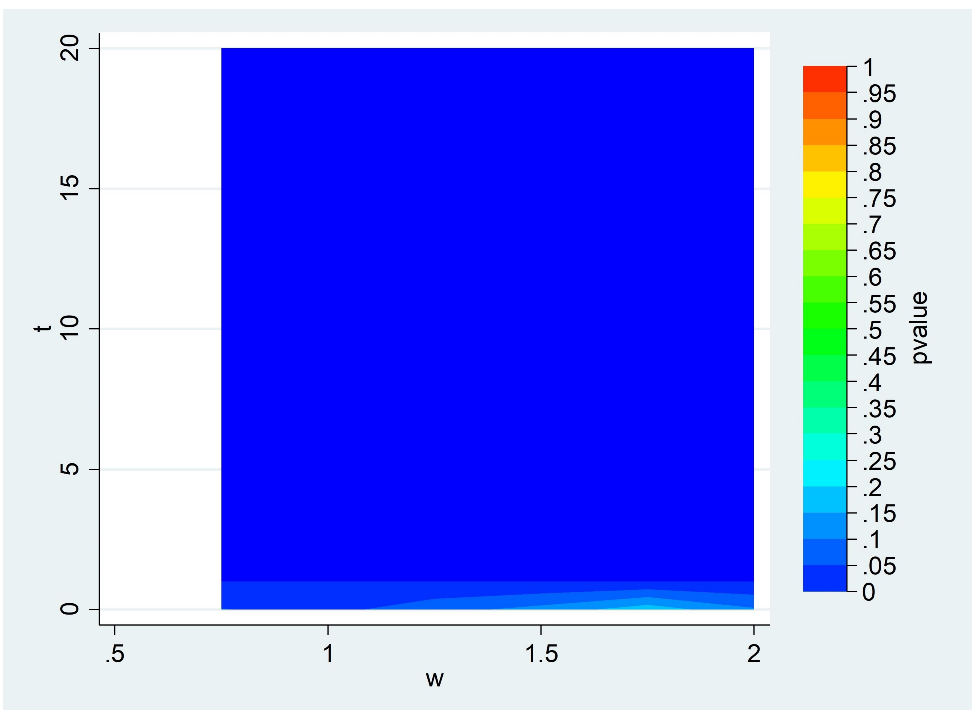


Figure 3.6: Sensitivity analysis for the RDD

The figure reports results from a sensitivity analysis under local randomization (see Cattaneo et al., 2016). We perform a sequence of hypotheses tests for different windows around the cutoff. Specifically, we show the test statistic of the null hypothesis of no treatment effect (x-axis) against the window length (y-axis). The p-values are calculated using randomization inference methods.



3.6 Tables

Table 3.1: Data and variable definitions

Variable	Description
A. Dimension of the data	
Individuals	Loan applicants who have an exclusive relationship with the bank and are majority owners (own more than 50%) of a firm. These borrowers apply to the bank for one or more business loans during the period 2002-2016 and the loan is either originated or denied. Due to the exclusive relationship, the bank holds information on the individuals' income and wealth even outside the year of loan application.
Year	The years covering the period 2002-2016.
B. Dependent variable	
Income	The euro amount of individuals' total annual income (in log).
C. Explanatory Variables: Running variable and cutoff	
Credit score	The credit score of the applicant, as calculated by the bank. We normalize this variable to take values around the cutoff of 0. The bank originates the loan if the credit score is higher than 0 and denies the loan otherwise.
Granted	A dummy variable equal to 1 if the loan is originated (Credit score>640) and 0 otherwise (Credit score<640).
D. Other covariates	
Education	An ordinal variable ranging between 0 and 5 if the individual completed the following education. 0: No secondary; 1: Secondary; 2: Post-secondary, non-tertiary; 3: Tertiary; 4: MSc, PhD or MBA.
Firm size	Total firm's assets (in log).
Firm leverage	The ratio of firm's total debt to total assets.
Gender	A dummy variable equal to 1 if the applicant is a male and 0 otherwise.
Loan amount	Log of the loan facility amount in thousands of euros.
Maturity	Loan duration in months.
Wealth	The euro amount of individuals' total wealth, as estimated by the bank (in log).
Initial wealth	Individuals' wealth in the first year before the loan application in which this information is available (one to five years before).

Table 3.2: Summary statistics

The table reports summary statistics (number of observations, mean, standard deviation, minimum, and maximum) for the variables used in the empirical analysis. The variables are defined in Table 1.

	Obs.	Mean	St. dev.	Min.	Max.
Income	61,863	11.01	0.376	9.852	12.29
Income t-1	57,682	10.58	0.406	9.804	12.62
Income t+1	57,766	11.1	0.388	9.866	12.58
Income t+3	49,514	11.14	0.373	9.987	12.57
Income t+5	41,391	11.16	0.363	10.04	12.62
Granted	61,863	0.867	0.498	0	1
Credit score	61,863	0.103	1.205	-2.921	2.1
Education	61,863	2.975	1.018	0	5
Firm size	61,863	12.821	0.806	2.5	12.03
Firm leverage	61,863	0.207	0.0249	0.143	0.917
Loan amount	61,863	2.323	0.845	0.679	7.48
Maturity	61,863	34.35	10.14	7	103
Wealth	61,863	12.14	0.556	8.564	14.05
Initial wealth	40,953	12.09	0.406	7.952	14.2

Table 3.3: Manipulation test

The table reports results from the manipulation testing procedure using the local polynomial density estimator proposed by Cattaneo et al. (2018). To perform this test, we rely on the local quadratic estimator with cubic bias-correction and triangular kernel. We report the conventional test statistic (Conventional) and the robust bias-corrected statistic (Robust) along with the corresponding p-value. The null hypothesis consists in no manipulation.

	T-stat	P-value
Conventional	1.5861	0.1127
Robust	1.2064	0.2277

Table 3.4: Results from parametric RDD

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is OLS on the RDD model of equation (2). Specifications (1) to (3) do not include any covariate besides the treatment and assignment variables. More covariates are included in specifications (4) to (6). The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Income t+1	Income t+3	Income t+5	Income t+1	Income t+3	Income t+5
Granted	0.0512***	0.0730***	0.0699***	0.0536***	0.0754***	0.0718***
	(0.0062)	(0.0064)	(0.0069)	(0.0063)	(0.0066)	(0.0072)
Credit score	-0.0015	0.006	0.0120***	-0.0056	0.0027	0.0084*
	(0.0038)	(0.0039)	(0.0042)	(0.0039)	(0.0041)	(0.0044)
Granted x Credit score	-0.0013	-0.0122**	-0.0216***	0.0026	-0.0087	-0.0168***
	(0.0052)	(0.0053)	(0.0057)	(0.0053)	(0.0056)	(0.0060)
Income t-1				0.0958***	0.0653***	0.0452***
				(0.0041)	(0.0043)	(0.0045)
Education				0.0023	-0.0017	0.0004
				(0.0016)	(0.0017)	(0.0019)
Firm size				-0.0004	0.003	-0.0015
				(0.0021)	(0.0022)	(0.0024)
Firm leverage				0.1872***	0.2877***	0.2435***
				(0.0672)	(0.0745)	(0.0778)
Loan amount				-0.0008	-0.0023	-0.0014
				(0.0020)	(0.0021)	(0.0023)
Maturity				0.0004**	0.0001	0.0002
				(0.0002)	(0.0002)	(0.0002)
Constant	11.0740***	11.1044***	11.1301***	9.9753***	10.3098***	10.5980***
	(0.0045)	(0.0047)	(0.0051)	(0.0517)	(0.0535)	(0.0558)
Observations	57,766	49,514	41,391	53,585	45,333	37,210
Adjusted R-squared	0.004	0.01	0.01	0.015	0.015	0.013
Clustering	Individual	Individual	Individual	Individual	Individual	Individual

Table 3.5: Results from non-parametric RDD

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator, the bias-corrected RD estimates with conventional variance estimator, and the bias-corrected RD estimates with robust variance estimator. Specifications (1) to (3) do not include any covariate besides the assignment variable (Credit score). More covariates are included in specifications (4) to (6). The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The bias-corrected RD estimator and the robust variance estimator are obtained according to Calonico et al. (2018) and Calonico at el. (2014), respectively.

	(1) Income t+1	(2) Income t+3	(3) Income t+5	(4) Income t+1	(5) Income t+3	(6) Income t+5
Conventional	0.0599***	0.0605***	0.107***	0.0623***	0.0605***	0.105***
	(0.0127)	(0.0134)	(0.0166)	(0.0126)	(0.0146)	(0.0170)
Bias-corrected	0.0632***	0.0572***	0.113***	0.0649***	0.0564***	0.112***
	(0.0127)	(0.0134)	(0.0166)	(0.0126)	(0.0146)	(0.0170)
Robust	0.0632***	0.0572***	0.113***	0.0649***	0.0564***	0.112***
	(0.0150)	(0.0159)	(0.0188)	(0.0150)	(0.0172)	(0.0194)
Obs.	57,766	49,514	41,391	53,585	45,333	37,210
Eff. obs. left of cutoff	8,731	7,510	4,487	8,274	6,171	4,061
Eff. obs. right of cutoff	9,186	7,855	4,686	8,670	6,398	4,232
BW estimate	61.37	61.3	44.03	62.61	54.76	44.08
BW bias	98.59	97	79.73	97.82	88.67	79.28

Table 3.6: Controlling for "initial" wealth

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. The table essentially replicates columns (3) to (6) of Table 5, the difference being the inclusion of Wealth t-5 as a control variable. Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator, the bias-corrected RD estimates with conventional variance estimator, and the bias-corrected RD estimates with robust variance estimator. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The bias-corrected RD estimator and the robust variance estimator are obtained according to Calonico et al. (2018) and Calonico at el. (2014), respectively.

	(1) Income t+1	(2) Income t+3	(3) Income t+5
Conventional	0.0646***	0.0491***	0.112***
	(0.0148)	(0.0171)	(0.0227)
Bias-corrected	0.0681***	0.0450***	0.121***
	(0.0148)	(0.0171)	(0.0227)
Robust	0.0681***	0.0450**	0.121***
	(0.0175)	(0.0202)	(0.0260)
Obs.	36,856	28,604	20,481
Eff. obs. left of cutoff	5,312	4,238	2,207
Eff. obs. right of cutoff	5,572	4,386	2,295
BW estimate	57.92	58.91	42.43
BW bias	91.65	94.75	74.35

Table 3.7: Controlling for the probability of loan application in the parametric RDD

The table reports coefficients and standard errors (in parentheses) from a two-stage treatment effects model estimated with maximum likelihood. The first stage models the probability that individuals apply for a loan in a given year (probit model). The second stage is equivalent to equation (2), but including the fitted value of *Instantaneous probability of loan application* (i.e., the hazard rate) obtained in the first stage. The dependent variable is given in the first row of the table and all variables are defined in Table 1. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1) Income t+1	(2) Income t+3	(3) Income t+5
Granted	0.0468***	0.0738***	0.0751***
	(0.0078)	(0.0082)	(0.0091)
Credit score	-0.0061	0.0053	0.0058
	(0.0048)	(0.0051)	(0.0057)
Granted x Credit score	0.0027	-0.01	-0.0137*
	(0.0065)	(0.0070)	(0.0076)
Instantaneous probability	0.0158***	0.0271***	0.0190***
of loan application	(0.0026)	(0.0027)	(0.0030)
		First-stage results Probability of application	
Income		0.060***	
		(0.0086)	
Wealth		0.0065***	
		(0.0016)	
Education		0.001	
		(0.0018)	
Firm size		0.004***	
		(0.0014)	
Firm leverage		0.273**	
		(0.1190)	
Gender		0.016**	
		(0.0073)	
Observations	34,448	28,662	23,049
Controls as in Table 4	Yes	Yes	Yes
Clustering	Individual	Individual	Individual

Table 3.8: Controlling for the probability of loan application in the non-parametric RDD

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. The table essentially replicates the analysis of columns 4-6 of Table 5, the difference being the inclusion of *Instantaneous probability of loan application* (i.e., the hazard rate) obtained in the first stage as a control variable in a non-parametric estimation of equation (2). Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator, the bias-corrected RD estimates with conventional variance estimator, and the bias-corrected RD estimates with robust variance estimator. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The bias-corrected RD estimator and the robust variance estimator are obtained according to Calonico et al. (2018) and Calonico at el. (2014), respectively.

	(1) Income t+1	(2) Income t+3	(3) Income t+5
Conventional	0.0388***	0.0449**	0.0934***
	(0.0142)	(0.0180)	(0.0207)
Bias-corrected	0.0415***	0.0445**	0.100***
	(0.0142)	(0.0180)	(0.0207)
Robust	0.0415**	0.0445**	0.100***
	(0.0169)	(0.0217)	(0.0241)
Obs.	34,448	28,662	23,049
Eff. obs. left of cutoff	6,136	3,909	2,738
Eff. obs. right of cutoff	6,400	4,031	2,852
BW estimate	71.87	54.16	47.4
BW bias	112.34	82.27	79.21

Table 3.9: Inequality measures

Panel A reports the Gini coefficient and the Theil index for individuals' income at time t and time $t + 5$ around the cutoff (credit score $< |0.1|$). Panel B compares the equivalent Gini coefficients and Theil indices for the samples of granted and non-granted loans.

	Income t	Income t+5
Panel A. Inequality measures around the cutoff		
Gini coefficient	0.207	0.226
Theil index	0.067	0.074
Panel B. Inequality measures for accepted vs. denied applicants		
Credit is granted		
Gini coefficient	0.224	0.200
Theil index	0.080	0.065
Credit is denied		
Gini coefficient	0.193	0.214
Theil index	0.058	0.073

Table 3.10: Heterogeneity due to applicants' location

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator, the bias-corrected RD estimates with conventional variance estimator, and the bias-corrected RD estimates with robust variance estimator. All specifications include the control variables as in Table 8. The first three and the last three specifications distinguish lower and higher income regions based on our sample's median. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The robust variance estimator is obtained according to Calonico at el. (2014).

	Low income			High income		
	(1)	(2)	(3)	(4)	(5)	(6)
	Income t+1	Income t+3	Income t+5	Income t+1	Income t+3	Income t+5
Robust	0.0642**	0.0710***	0.1203***	0.0605***	0.0597**	0.0926***
	(0.0279)	(0.0230)	(0.0380)	(0.0191)	(0.0182)	(0.0263)
Obs.	28,883	24,757	20,696	28,883	24,757	20,695
Eff. obs. left of cutoff	4,220	3,412	2,311	4,113	3,347	2,290
Eff. obs. right of cutoff	4,355	3,504	2,384	4,160	3,416	2,297
BW estimate	58.6	56.28	43.28	55.69	55.11	41.18
BW bias	94.3	88.25	75.61	92.5	88.26	72.16

Table 3.11: Pre-post crisis

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator, the bias-corrected RD estimates with conventional variance estimator, and the bias-corrected RD estimates with robust variance estimator. All specifications include the control variables as in specifications 4-6 in Table 4. Specifications 1 to 3 are estimated using loan applications for the years 2009-2016 and specifications 4-6 using loan applications for 2000-2008 The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The bias-corrected RD estimator and the robust variance estimator are obtained according to Calonico et al. (2018) and Calonico at el. (2014), respectively.

	Crisis and post-crisis			Pre-crisis		
	(1)	(2)	(3)	(4)	(5)	(6)
	Income t+1	Income t+3	Income t+5	Income t+1	Income t+3	Income t+5
Conventional	0.0552***	0.0722***	0.105***	0.0627***	0.0456**	0.0964***
	(0.0211)	(0.0215)	(0.0201)	(0.0144)	(0.0178)	(0.0249)
Bias-corrected	0.0610***	0.0700***	0.112***	0.0639***	0.0395**	0.104***
	(0.0211)	(0.0215)	(0.0201)	(0.0144)	(0.0178)	(0.0249)
Robust	0.0610**	0.0700***	0.112***	0.0639***	0.0395*	0.104***
	(0.0249)	(0.0258)	(0.0229)	(0.0172)	(0.0207)	(0.0291)
Obs.	20,850	20,850	20,850	32,735	24,483	16,360
Eff. obs. left of cutoff	3,509	2,977	2,992	5,613	3,886	1,778
Eff. obs. right of cutoff	3,657	3,099	3,110	5,876	4,040	1,874
BW estimate	68.69	58.09	58.34	69.29	63.39	43.29
BW bias	109.9	87.97	103.87	106.17	108.54	72.05

3.7 Appendix

The Appendix reports results from additional sensitivity tests. In Table A3.1 we include several fixed effects in the parametric model. In Table A3.2 we use different bandwidth-selection rules. In Table A3.3 we include *Initial wealth* in the parametric RDD and Figure A3.1 illustrates *Initial wealth* around the cutoff.

Figure A3.1: Initial wealth around the cutoff

The figure reports Wealth t-5 (first instance of wealth before the loan application) against the Credit score. The continuous line represents a fourth order polynomial fit used to approximate the conditional mean of each covariate below and above the cutoff.

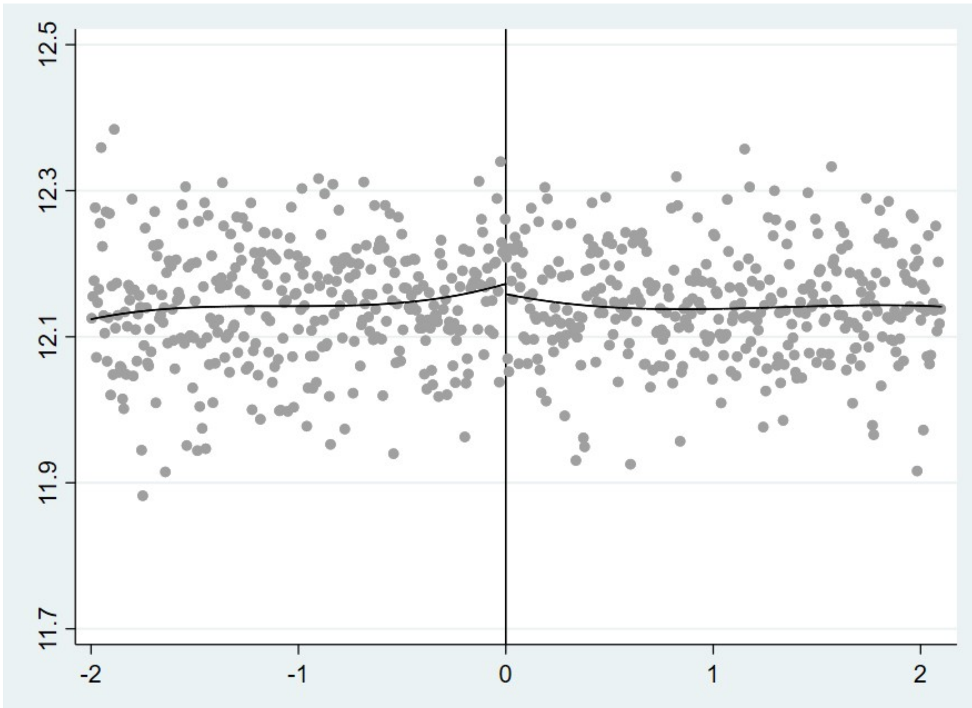


Table A3.1: Including industry, loan type, and year fixed effects in the parametric RDD

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is OLS on the RDD model of equation (2). Specifications (1) to (3) do not include any covariate besides the treatment and assignment variables. More covariates are included in specifications (4) to (6). All specifications include industry, loan type, and year fixed effects. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1) Income t+1	(2) Income t+3	(3) Income t+5	(4) Income t+1	(5) Income t+3	(6) Income t+5
Granted	0.0534***	0.0751***	0.0713***	0.0536***	0.0754***	0.0718***
	(0.0063)	(0.0066)	(0.0072)	(0.0063)	(0.0066)	(0.0072)
Credit score	-0.0051	0.0029	0.0089**	-0.0056	0.0027	0.0084*
	(0.0038)	(0.0040)	(0.0044)	(0.0039)	(0.0041)	(0.0044)
Granted x Credit score	0.0021	-0.0089	-0.0172***	0.0025	-0.0087	-0.0168***
	(0.0052)	(0.0055)	(0.0059)	(0.0053)	(0.0056)	(0.0060)
Income t-1				0.0975***	0.0657***	0.0447***
				(0.0053)	(0.0056)	(0.0058)
Education				0.0023	-0.0017	0.0004
				(0.0016)	(0.0017)	(0.0019)
Firm size				-0.0004	0.003	-0.0015
				(0.0021)	(0.0022)	(0.0024)
Firm leverage				0.1872***	0.2877***	0.2435***
				(0.0672)	(0.0745)	(0.0778)
Loan amount				-0.0008	-0.0023	-0.0014
				(0.0020)	(0.0021)	(0.0023)
Maturity				0.0004**	0.0001	0.0002
				(0.0002)	(0.0002)	(0.0002)
Constant	0.0429***	0.0297***	0.0209***	-0.002	-0.0004	0.0005
	(0.0029)	(0.0030)	(0.0032)	(0.0038)	(0.0039)	(0.0041)
Observations	53,585	45,333	37,210	53,585	45,333	37,210
Clustering	Individual	Individual	Individual	Individual	Individual	Individual

Table A3.2: Alternative bandwidth selection methods

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator. The specifications do not include any covariate besides the assignment variable (credit score). Specifications (1), (3), and (5) use the two mean squared error (MSE)-optimal bandwidth selectors (below and above the cutoff) for the RD treatment effect. Specifications (2), (4), and (6) use one common coverage error (CER)-optimal bandwidth selector for the RD treatment effect. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias.

	(1) Income t+1	(2) Income t+1	(3) Income t+3	(4) Income t+3	(5) Income t+5	(6) Income t+5
	0.0611***	0.0716***	0.0610***	0.0645***	0.103***	0.0956***
	(0.0127)	(0.0167)	(0.0131)	(0.0178)	(0.0159)	(0.0215)
Obs.	57,766	57,766	49,514	49,514	41,391	41,391
Eff. obs. left of cutoff	7,743	5,053	8,260	4,373	5,180	2,599
Eff. obs. right of cutoff	10,530	5,284	7,802	4,536	4,831	2,738

Table A3.3: Controlling for "initial" wealth: OLS results

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is OLS on the RDD model of equation (2). The table essentially replicates columns (3) to (6) of Table 4, the difference being the inclusion of Wealth t-5 as a control variable. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1) Income t+1	(2) Income t+3	(3) Income t+5
Granted	0.0514***	0.0726***	0.0814***
	(0.0072)	(0.0080)	(0.0094)
Credit score	-0.0071	-0.0023	0.0003
	(0.0044)	(0.0050)	(0.0059)
Granted x Credit score	0.0028	-0.002	-0.0083
	(0.0060)	(0.0068)	(0.0079)
Income t-1	0.0816***	0.0600***	0.0450***
	(0.0051)	(0.0056)	(0.0064)
Education	0.0032*	-0.0027	0.0013
	(0.0018)	(0.0021)	(0.0024)
Firm size	-0.0001	0.0024	-0.0007
	(0.0024)	(0.0027)	(0.0031)
Firm leverage	0.1898**	0.1764**	0.2908***
	(0.0765)	(0.0850)	(0.1051)
Loan amount	0.0001	0.0014	0.0006
	(0.0023)	(0.0026)	(0.0030)
Maturity	0.0004*	0	0.0001
	(0.0002)	(0.0002)	(0.0003)
Wealth t-5	0.0215***	0.0148***	0.0046
	(0.0032)	(0.0035)	(0.0040)
Constant	9.9057***	10.2427***	10.5395***
	(0.0736)	(0.0803)	(0.0929)
Observations	36,856	28,604	20,481
Clustering	Individual	Individual	Individual

References

Acharya, V., Eisert, T., Eufinger, C., Hirsch, C., Forthcoming. Real effects of the sovereign debt crisis in Europe: Evidence from syndicated loans. *Review of Financial Studies.*

Adams, R. B., Ferreira, D., 2009. Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics* 94, 291-309.

Adams, W., Einav, L., Levin, J., 2009. Liquidity constraints and imperfect information in subprime lending. *American Economic Review* 99 (1), 49-84.

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., Evanoff, D. D., 2010. Learning to cope: Voluntary financial education and loan performance during housing crisis. *American Economic Review: Papers and Proceedings* 100, 495-500.

Allen, F., Carletti, E., Marquez, R., 2011. Credit market competition and capital regulation. *Review of Financial Studies* 24 (4), 983-1018.

Amiti, M., Weinstein, D. E., Forthcoming. How much do idiosyncratic bank shocks affect investment? Evidence from matched bank-firm loan data. *Journal of Political Economy.*

Andrews, I., Stock, J., Sun, L., 2018. Weak instruments in IV regression: Theory and practice. Working paper.

Angrist, J. D., Pischke, J. S., 2009. Mostly harmless econometrics. Princeton University Press.

Badarinza, C., Campbell, J. Y., Ramadorai, T., 2018. What calls to arms? International evidence on interest rates and the choice of adjustable-rate mortgages. *Management Science* 64 (5), 2275-2288.

Balduzzi, P., Brancati, E., Schiantarelli, F., 2017. Financial markets, banks' cost of funding, and firms' decisions: lessons from two crises. *Journal of Financial Intermediation* 36, 1-15.

Banerjee, A. V., Newman, A. F., 1993. Occupational choice and the process of development. *Journal of Political Economy* 101 (2), 274-298.

Basten, C., Guin, B., Koch, C., 2017. How do banks and households manage interest rate risk? Evidence from the Swiss mortgage market. Cesifo Working Paper No. 6649.

BCBS, 2012. Core principels for effective banking supervision. Technical report.

Beck, T., Levine, T., Levkov, A., 2010. Big bad banks? The winners and losers from bank deregulation in the United States. *Journal of Finance* 65 (5), 1637-1667.

Becker, S. O., Egger, P. H., Von Ehrlich, M., 2013. Absorptive capacity and the growth and investment effects of regional transfers: A regression discontinuity design with heterogeneous treatment effects. *American Economic Journal: Economic Policy* 5 (4), 29-77.

Bentolila, S., Jansen, M., Jiminéz, G., Ruano, S., 2017. When credit dries up: Job losses in the great recession. *Journal of the European Economic Association* 16 (3), 650-695.

Berg, T., 2018. Got rejected? Real effects of not getting a loan. *Review of Financial Studies* 31 (12), 4912-4957.

Bhat, G., Desai, H., 2017. Bank capital and loan monitoring. Working Paper.

Bond, S., Ham, K. Y., Maffini, G., Nobili, A., Ricotti, G., 2016. Regulation, tax and capital structure: evidence from administrative data on Italian banks. Bank of Italy Occasional Paper N. 361.

Brunner, A., Krahnen, J. P., 2008. Multiple lenders and corporate distress: Evidence on debt restructuring. *Review of Economic Studies* 75, 415-442.

Brunnermeier, M. K., 2009. Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic Perspectives* 23 (1), 77-100.

Calonico, S., Cattaneo, M. D., Titiunik, R., 2014. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica* 82 (6), 2295-2326.

Calonico, S., Cattaneo, M. D., Farrell, M. H., 2018. On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association* 0 (0), 1-13.

Cameron, A. C., Gelbach, J. B., Miller, D. L., 2011. Robust inference with multiway clustering. *Journal of Business and Economic Statistics* 29 (2), 238-249.

Cameron, A. C., Miller, D. L., 2015. A practitioner's guide to cluster-robust inference. *Journal of Human Resources* 50 (2), 317-372.

Campbell, J. Y., 2012. Mortgage market design. *Review of Finance* 17, 1-33.

Campbell, J. Y., Cocco, J. F., 2003. Household risk management and optimal mortgage choice. The Quarterly *Journal of Economics* 118 (4), 1449-1494.

Carletti, E., De Marco, F., Ioannidou, V., Sette, E., 2018. Banks as patient lenders: Evidence from a tax reform. Working paper.

Carvalho, I., 2017. What is the impact of the access to external finance on the capital structure of SMEs in Europe? PhD Dissertation, Catòlica Lisbon School of Business & Economics.

Cattaneo, M. D., Titiunik, R., Vazquez-Bare, G., 2016. Inference in regression discontinuity design under local randomization. *Stata Journal* 16 (2), 331-367.

Cattaneo, M. D., Jansson, M., Ma, X., 2018. Manipulation testing based on density discontinuity. *Stata Journal* 18 (1), 234-261.

Célérier, C., Kick, T., Ongena, S., 2018. Taxing bank leverage: The effects on bank capital structure, credit supply and risk-taking. Working Paper.

Cerqueiro, G., Ongena, S., Roszbach, K., 2016. Collateralization, bank loan rates, and monitoring. *Journal of Finance* 71 (3), 1295-1321.

Cetorelli, N., Goldberg, L. S., 2011. Global banks and international shock transmission: Evidence from the crisis. *IMF Economic Review* 59 (1), 41-76.

Charness, G., Gneezy, U., 2012. Strong evidence of gender differences in risk taking. *Journal of Economic Behavior and Organization* 83, 50-58.

Chodorow-Reich, G., 2014. The employment effects of credit market disruptions: Firm-level evidence from the 2008-2009 financial crisis. *Quarterly Journal of Economics* 129, 1–59.

Cingano, F., Manaresi, F., Sette, E., 2013. Does credit crunch investments down? *Review of Financial Studies* 29, 2737–2773.

Clarke, G. R. G., Xu, L. C., Zou, H. F., 2006. Finance and income inequality: What do data tell us? *Southern Economic Journal* 72 (3), 578-596.

Cornelissen, T., Dustmann, C., Shönberg, U., 2016. From LATE to MTE: Alternative methods for the evaluation of policy interventions. *Labour Economics* 41, 47-60.

Dass, N., Massa, M., 2011. The impact of a strong bank-firm relationship on the borrowing firm. *Review of Financial Studies* 24, 1204-1260.

De Haan, J., Sturm, J. E., 2017. Finance and income inequality: A review and new evidence. *European Journal of Political Economy* 50, 171-195.

Delis, M. D., Hasan, I., Kazakis, P., 2014. Bank regulations and income inequality: Empirical evidence. *Review of Finance* 18, 1811-1846.

Dell'Ariccia, G., Laeven, L., Marquez, R., 2014. Real interest rates, leverage, and bank risk-taking. *Journal of Economic Theory* 149, 65-99.

Dell'Ariccia, G., Laeven, L., Suarez, G. A., 2017. Bank leverage and monetary policy's risk-taking channel: Evidence from the United States. *Journal of Finance* 72 (2), 613-654.

Denk, O., Cournède, B., 2015. Finance and income inequality in OECD countries. OECD Economic Department Working Papers No. 1224.

Devereux, M. P., Vella, J., Johannesen, N., 2015. Can taxes tame the banks? Evidence from the European bank levies. Saïd Business School Research Papers.

Diamond, D.W., 1984. Financial intermediation and delegated monitoring. *Review of Economic Studies* 51 (3), 393-414.

Diamond, D.W., Rajan, R. G., 2000. A theory of bank capital. *Journal of Finance* 55 (6), 2431-2465.

Di Maggio, M., Kermani, A., Keys, B. J., Piskorski, T., Ramcharan, R., Seru, A., Yao, V., 2017. Interest rate pass-through: Mortgage rates, household consumption, and voluntary deleveraging. *American Economic Review* 107 (11), 3550-3588.

Duchin, R., Ozbas, O., Sensoy, B. A., 2010. Costly external finance, corporate investment, and the subprime mortgage crisis. *Journal of Financial Economics* 97, 418–35.

ECB, 2008. Covered bonds in the EU financial system. Technical report.

ECB, 2009. Housing finance in the euro area. Technical report. ECB Occasional Paper Series n. 101.

ECB, 2014. Guide to banking supervision. Technical report.

ECB, 2016. Economic bulletin. Technical report. Issue 6.

ECB, 2016. The household finance and consumption survey: Results from the second wave. ECB Statistics Paper No 18. December 2016. Household Finance and Consumption Network.

ECB, 2017. The ECB survey of professional forecasters. Technical report. Third quarter of 2017.

ECBC, 2016. 2016 ECBC European covered bond fact book. Technical report.

Eckel, C. C., Grossman, P. J., 2002. Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior* 23 (4), 281-295.

Eckel, C. C., Grossman, P. J., 2008. Men, women and risk aversion: Experimental evidence. Handbook of Experimental Economics Results, Volume I, Chapter 113, Elsevier.

Ehrmann, M., Ziegelmeyer, M., 2017. Mortgage choice in the euro area: Macroeconomic determinants and the effect of monetary policy. *Journal of Money, Credit and Banking* 49 (2-3), 469-494.

Foà, G., Gambacorta, L., Guiso, L., Mistrulli, P. E., 2015. The supply side of household finance. Technical report. Bank of Italy Working Papers n. 1044.

Fornero, E., Monticone, C., Trucchi, S., 2011. The effect of financial literacy on mortgage choice. Technical report. CeRP Working Paper n. 121/11.

Freixas, X., Rochet, J. C., 2008. Microeconomics of banking. Second edition. The MIT Press.

Fuster, A., Vickery, J., 2014. Securitization and the fixed-rate mortgage. *Review of Financial Studies* 28 (1), 176-211.

Gale, D., Hellwig, M., 1985. Incentive-compatible debt contracts: The one-period problem. *Review of Economic Studies* 52, 647-663.

Galor, O., Zeira, J., 1993. Income distribution and macroeconomics. *Review of Economic Studies* 60 (1), 35-52.

Gambacorta, L., Ricotti, G., Sundaresan, S., Wang, Z., 2017. The effects of tax on bank liability structure. BIS Working Papers N. 611.

Gambacorta, L., van Rixtel, A., Schiaffi, S., 2017. Changing business models in international bank funding. BIS Working Papers No 614.

Gan, J., 2007. The real effects of asset market bubbles: Loan- and firm-level evidence of a lending channel. *Review of Financial Studies* 20, 1941–73.

Gathergood, J., Weber, J., 2017. Financial literacy, present bias and alternative mortgage products. *Journal of Banking and Finance* 78, 58-83.

Gertler, M., Gilchrist., S., 1994. Monetary policy, business cycles, and the behavior of small manufacturing firms. *Quarterly Journal of Economics* 109, 309–40.

Greene, W., 2004a. The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal* 7 (1), 98-119.

Greene, W., 2004b. Fixed effects and bias due to the incidental parameters problem in the Tobit model. *Econometric Reviews* 23 (2), 125-147.

Greenstone, M., Mas, A., Nguyen, H. L., Forthcoming. Do credit market shocks affect the real economy? Quasi-experimental evidence from the great recession and "normal" economic times. *American Economic Journal: Economic Policy.*

Greenwood, J., Jovanovic, B., 1990. Financial development, growth, and the distribution of income. *Journal of Political Economy* 98 (5), 1076-1107.

Gu, W. G., de Mooiji, R., Poghsyan, T., 2015. Taxation and leverage in international banking. *International Tax and Public Finance* 22, 177-200.

Guren, A. M., Krishnamurthy, A., McQuade, T. J., 2018. Mortgage design in an equilibrium model of the housing market. NBER Working Paper No. 24446.

Gustafson, M. T., Ivanov, I. T., Meisenzahl, R. R., 2017. Bank monitoring: Evidence from syndicated loans. Working paper.

Hamori, S., Haschiguchi, Y., 2012. The effect of financial deepening on inequality: Some international evidence. *Journal of Asian Economics* 23, 353-359.

Heckman, J. J., 1976. The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement* 5, 475–492.

Heckman, J. J., Urzuna, S., Vytlacil, E., 2006. Understanding instrumental variables in models with essential heterogeneity. IZA Discussion Paper No. 2320.

Hoffmann, P., Langfield, S., Pierobon, F., Vuillemey, G., 2017. Who bears interest rate risk? Working Paper.

Holmstrom, B., Tirole, J., 1997. Financial intermediation, loanable funds, and the real sector. *Quarterly Journal of Economics* 112 (3), 663-691.

Holt, C. A., Laury, S. K., 2002. Risk aversion and incentive effects. *American Economic Review* 92 (5), 1644-1655.

Houston, J., James, C., Marcus, D., 1997. Capital market frictions and the role of internal capital markets in banking. *Journal of Financial Economics* 46, 135-164.

Ippolito, F., Ozdagli, A. K., Perez-Orive, A., 2017. The transmission of monetary policy through bank lending: The floating rate channel. Technical report. Finance and Economics Discussion Series 2017-026. Washington: Board of Governors of the Federal Reserve System.

Imbens, G. W., Lemieux, T., 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142, 615-635.

IMF, 2009. World economic outlook. Crisis and recovery. April 2009. World Economic and Financial Surveys.

Iyer, R., Peydró, J. L., da-Rocha-Lopes, S., Schoar, A., 2014. Interbank liquidity crunch and the firm credit crunch: evidence from the 2007-2009 crisis. *Review of Financial Studies* 27 (1), 347-372.

Iyer, R., Puri, M., 2012. Understanding bank runs: The importance of depositor-bank relationships and networks. *American Economic Review* 102 (4), 1414-1445.

Jauch, S., Watzka, S., 2016. Financial development and income inequality: A panel data approach. *Empirical Economics* 51, 291-314.

Jianakoplos, N. A., Bernasek, A., 1998. Are women more risk averse? *Economic Inquiry* 36, 620-630.

Jiménez, G., Ongena, S., Peydró, J. L., Saurina, J., 2012. Credit supply and monetary policy: Identifying the bank balance-sheet channel with loan applications. *American Economic Review* 102, 2301–2326.

Jiménez, G., Ongena, S., Peydró, J. L., Saurina, J., 2014. Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking? *Econometrica* 82 (2), 463-505.

Jiménez, G., Ongena, S., Peydró, J. L., Saurina, J., 2017. Macroprudential policy, countercyclical bank capital buffers and credit supply: Evidence from the Spanish dynamic provisioning experiments. *Journal of Political Economy* 125 (6), 2126-2177.

Kappel, V., (2010). The effects of financial development on income inequality and poverty. CER-ETH Working Paper 10/127.

Kashyap, A., Stein., J., 2000. What do a million observations on banks say about the transmission of monetary policy. *American Economic Review* 90, 407–28.

Keen, M., de Mooji, R. A., 2016. Debt, taxes and banks. *Journal of Money, Credit and Banking* 48 (1), 5-33.

Kirti, D., 2017. Why do bank-dependent firms bear interest-rate risk? Technical report. IMF Working Paper n. 17/3.

Klein, M. W., Peek, J., Rosengren, E. S., 2002. Troubled banks, impaired foreign direct investment: The role of relative access to credit. *American Economic Review* 93 (3), 664-682.

Koijen, R. S., Hemert, O. V., Nieuwerburgh, S. V., 2009. Mortgage timing. *Journal of Financial Economics* 93, 292-324.

Krasa, C. M., Villamil, A. P., 1992. Monitoring the monitor: An incentive structure for a financial intermediary. *Journal of Economic Theory* 57, 197-221.

Kruse, J. B., Thompson, M. A., 2003. Valuing low probability risk: Survey and experimental evidence. *Journal of Economic Behavior and Organization* 50, 495-505.

Kumhof, M., rancière, R., 2010. Inequality, leverage and crises. IMF Working Paper 10/268.

Lancaster, T., 2000. The incidental parameters problem since 1948. *Journal of Econometrics* 95 (2), 391-413.

Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. *Journal of Economic Literature* 48, 281-355.

Li, H., Squire, L., Zou, H. F., 1998. Explaining international and intertemporal variations in income inequality. *Economic Journal* 108, 26-43.

MEF, 2012. Relazione generale sulla situazione economica del paese 2012.

Mehran, H., Thakor, A., 2011. Bank capital and value in the cross-section. *Review of Financial Studies* 24 (4), 1019-1067.

Moore, E., Eckel, C. C., 2003. Measuring ambiguity aversion. Working Paper.

Naceur, S. B., Zhang, R., 2016. Financial development, inequality and poverty: Some international evidence. IMF Working Paper 16/32.

Neyman, J., Scott, E. L., 1948. Consistent estimates based on partially consistent observations. *Econometrica* 16 (1), 1-32.

OECD, 2015. In it together: Why less inequality benefits all. Overview of inequality trends, key findings and policy directions. OECD Publishing.

Ongena, S., Smith, D. C., 2000. What determines the number of bank relationships? Cross-country evidence. *Journal of Financial Intermediation* 9, 26-56.

Paiella, M., Pozzolo, A. F., 2007. Household Credit Usage, chapter Choosing Between Fixed and Adjustable Rate Mortgages, 219-235. Palgrave Macmillan.

Petersen, M. A., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1), 435-480.

Plosser, M. C., Santos, J. A. C., 2016. Bank monitoring. Working paper.

Popov, A., Rocholl., J., Forthcoming. Financing constraints, employment, and labor compensation. *Journal of Financial Intermediation.*

Rajan, R. G., 1992. Insiders and outsiders: The choice between informed and arm's-length debt. *Journal of Finance* 47 (4), 1367-1400.

Rajan, R. G., 2010. Fault lines. How hidden fractures still threaten the world economy. Princeton University Press.

Schepens, G., 2016. Taxes and bank capital structure. *Journal of Financial Economics* 120, 585-600.

Schnabl, P., 2012. The international transmission of bank liquidity shocks: Evidence from an emerging market. *Journal of Finance* 67 (3), 897-932.

Schubert, R., Brown, M., Gysler, M., Brachinger, H. W., 1999. Gender and economic transactions. *AEA Papers and Proceedings* 89 (2), 381-385.

Sharpe, S., 1990. Asymmetric information, bank lending and implicit contracts: A stylized model of consumer relationships. *Journal of Finance* 45 (4), 1069-1087.

Solt, Frederick. 2016. “The Standardized World Income Inequality Database”. *Social Science Quarterly* 97. SWIID Version 6.2, March 2018.

Staiger, D., Stock, J., 1997. Instrumental variables regression with weak instruments. *Econometrica* 65, 557-586.

Stanga, I., Vlahu, R., de Haan, J., 2017. Mortgage arrears, regulation and institutions: Cross-country evidence. DNB Working Paper n. 580.

Stock, J. H, Yogo, M., 2005. Identification and inference for econometric models: Essays in honor of Thomas Rothenberg. Chapter *Testing for Weak Instruments in Linear IV Regression*, Cambridge University Press, 80-108.

Townsend, R., 1979. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory* 21 (20), 265-293.

Wooldridge, J. M., 2010. Econometric Analysis of Cross Section and Panel Data. MIT Press.

Curriculum Vitae

Personal Details

Name	Fulvia Fringuellotti
Date of Birth	December 1, 1988
Place of Birth	Terni, Italy
Nationality	Italian

Education

09/2013-07/2019	**Swiss Finance Institute PhD Program in Finance** University of Zurich, Department of Banking and Finance (Zurich, Switzerland) Supervisor: Prof. Dr. Steven Ongena
09/2010-07/2012	**Master's Degree in Economics and Finance** LUISS Guido Carli University (Rome, Italy)
09/2007-07/2010	**Bachelor's Degree in Economics, Financial Markets and Intermediaries** LUISS Guido Carli University (Rome, Italy)

Professional Experience

09/2014-07/2019	Research and Teaching Assistant at University of Zurich, Chair of Corporate Finance, Prof. Dr. Michel Habib (Zurich, Switzerland)
09/2012-08/2013	Consultant at Be Think, Solve, Execute S.p.A. (Milan, Italy)